

Vaccine production and glycoengineering in *Arabidopsis thaliana* seeds

Robin Piron

April 2014

Promoters: Prof. Dr. Ann Depicker¹

Prof. Dr. Nico Callewaert²

Ghent University, Faculty of Sciences:

¹Department of Plant Biotechnology and Bioinformatics

²Department of Biochemistry and Microbiology

¹VIB – Department of Plant Systems Biology

²VIB – Inflammation Research Center

Thesis submitted in partial fulfillment of the requirements for the degree of
Doctor (PhD) in Sciences: Biotechnology



This research was performed in the Plant-made Antibodies and Immunogens research group led by Prof. Dr. Ann Depicker at the VIB-UGent Department of Plant Systems Biology.

A part of the work was conducted in collaboration with the group of Prof. Dr. Nico Callewaert at the VIB-UGent Inflammation Research Center, that of Prof. Dr. Hans Nauwynck at the UGent Department of Virology, Parasitology and Immunology and that of Prof. Dr. Johan Grooten at the UGent Department of Biomedical Molecular Biology.



The research was funded by a PhD grant for Strategic Basic Research from the agency of Innovation by Science and Technology.

Contents

Examination Board.....	viii
Abbreviations.....	ix
Scope & objectives.....	1
Summary.....	3
Samenvatting.....	7
Chapter I	11
Towards a viral subunit vaccine produced in <i>A. thaliana</i> seeds, an introduction.....	11
Vaccination.....	13
The early days.....	13
Attenuated, inactivated and subunit vaccines	13
Recombinant vaccines.....	14
Transgenic plants as biofactories.....	15
Seed production of PMPs.....	18
The <i>Arabidopsis</i> seed platform.....	19
The porcine reproductive and respiratory syndrome virus.....	21
Chapter II	27
Towards a viral subunit vaccine produced in <i>A. thaliana</i> seeds, results: “Boosting in planta production of antigens derived from the porcine reproductive and respiratory syndrome virus (PRRSV) and subsequent evaluation of their immunogenicity”.	27
Abstract.....	29
Introduction.....	29
Results.....	32
Antigen expression in <i>A. thaliana</i> seed	32
Glycosylation of antigens produced in <i>A. thaliana</i> seed	37
Accumulation levels of antigens produced in <i>A. thaliana</i> seed	40
Purification of antigens produced in <i>A. thaliana</i> seed	41
Immunogenicity of antigens produced in <i>A. thaliana</i> seeds	43

Discussion.....	47
Material and methods.....	51
Ethics Statement	51
Cloning and generation of transgenic lines	51
Protein extraction, protein concentration determination, SDS-PAGE, Coomassie staining and western blot analysis.....	53
Glycan analysis.....	54
Antigen purification by IMAC.....	54
Antigen purification by protein-A affinity chromatography.....	55
Antigen purification by gel filtration	55
ELISA to determine antigen accumulation in the seed	55
Immunization of mice	56
Characterization of the antibody response in mice	57
Immunization of piglets.....	57
Characterization of the antibody response in piglets	58
Acknowledgments.....	58
Supplemental information	59
Chapter III	65
Glycoengineering in <i>Arabidopsis</i> seeds, an introduction.....	65
Glycosylation	67
Plant N-glycosylation.....	68
Immunogenicity of plant N-glycans	72
‘Humanizing’ the plant N-glycosylation pathway	74
A key for each lock, towards the desired glycoform.....	78
Engineering the <i>Arabidopsis</i> seed platform, Glycodelete Technology™	80
Rationale.....	80
The <i>Trichoderma reesei</i> Endo-N-acetyl-β-D-glucosaminidase	81
The <i>Arabidopsis thaliana</i> complex-glycan-deficient mutant	83

Chapter IV	85
Glycoengineering in <i>Arabidopsis</i> seeds, results & discussion.	85
A reporter protein for the GlycoDelete technology, the activation-associated secretory protein 1	87
Results	89
EndoT expression in <i>Arabidopsis thaliana</i> complex-glycan-deficient seeds	89
EndoT deglycosylates endogenous seed glycoproteins	93
ASP1 expression in wild type, cgl and glycodelete™ <i>Arabidopsis</i> seed	95
EndoT accumulation level and stability over generations.....	99
Purification of recombinant ASP1.....	100
Glycan analysis of ASP1 produced in wild type, cgl and GlycoDelete <i>Arabidopsis</i> seeds.....	101
ASP1 carrying complex plant N-glycans, but not GlycoDelete ASP1 induces an anti-carbohydrate response in rabbits.....	110
Discussion.....	113
Material and methods.....	118
Cloning and generation of transgenic lines	118
Protein extraction, protein concentration determination, SDS-PAGE, Coomassie staining and western blot analysis.....	119
PNGaseF digestion of recombinant endoT	120
ELISA to determine endoT accumulation level.....	120
Recombinant ASP1 purification	121
Recombinant ASP1 glycan analysis by endoT digest.....	121
Recombinant ASP1 glycan analysis by glycosidase treatment and DSA-FACE	122
Recombinant ASP1 glycan analysis by MALDI-TOF spectrometry	122
Immunization of rabbits	123
Characterization of the antibody response in rabbits.....	124
Supplemental information	126
Chapter V	135
Conclusions & perspectives.....	135

Vaccine production in <i>Arabidopsis</i> seeds.....	137
Glycoengineering <i>Arabidopsis</i> seeds.....	140
References	144
Addendum	163
Acknowledgements.....	163
Curriculum Vitae	165
Publications	167

Examination Board

Chair:

Prof. Dr. Geert De Jaeger UGent, Faculty of Sciences
Department of Plant Biotechnology and Bioinformatics
VIB, Department of Plant Systems Biology

Promoters:

Prof. Dr. Ann Depicker UGent, Faculty of Sciences
Department of Plant Biotechnology and Bioinformatics
VIB, Department of Plant Systems Biology

Prof. Dr. Nico Callewaert UGent, Faculty of Sciences
Department of Biochemistry and Microbiology
VIB, Inflammation Research Center

Examination commission:

Prof. Dr. Els Van Damme* UGent, Faculty of Bioscience Engineering
Department of Molecular Biotechnology

Prof. Dr. Peter Geldhof* UGent, Faculty of Veterinary Medicine
Department of Virology, Parasitology and immunology

Prof. Dr. Geert Angenon* VUB, Faculty of Sciences
Department of Applied Biological Sciences,

Dr. Stefaan De Koker* UGent, Faculty of Sciences
Department of Biomedical Molecular Biology

Dr. Wander Van Breedam* UGent, Faculty of Veterinary Medicine
Department of Virology, Parasitology and immunology

*member of the reading committee

Abbreviations

3'Arc	3' end of the arceline5-I gene
3'ocs	3' end of the octopine synthase gene
3'-UTR	3'-untranslated region
5'-UTR	5'-untranslated region
35S	Constitutive promoter of the cauliflower mosaic virus
Ab	Antibody
ADCC	Antibody dependent cellular cytotoxicity
Apo	Apoplast
APTS	8-amino-1,3,6-pyrenetrisulfonic acid
ASP1, 2	Activation-associated secreted protein 1 and 2
A1,3-fucT	Core α 1,3-fucosyltransferase
A1,4-fucT	α 1,4-fucosyltransferase
<i>Bar</i>	Bar gene conferring phosphinothricin resistance
Bp	Basepairs
BSA	Bovine serum albumin
B1,3-galT	β 1,3-galactosyltransferase
B1,4-galT	β 1,4-galactosyltransferase
CCD	Cross-reactive carbohydrate determinant
CDC	Complement dependent cytotoxicity
<i>Cgl1-1</i>	<i>Arabidopsis thaliana</i> complex-glycan-deficient 1-1 mutant
<i>Cgl1-2</i>	<i>Arabidopsis thaliana</i> complex-glycan-deficient 1-2 mutant
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate
CHCA	α -cyano-4-hydroxycinnamic acid
Col-0	<i>Arabidopsis thaliana</i> ecotype Columbia
DSA-FACE	DNA sequencer-assisted fluorofore-assisted carbohydrate electrophoresis
DNA	Deoxynucleic acid
EDTA	Ethylenediaminetetraacetic acid

EndoT	<i>Trichoderma reesei</i> endo- <i>N</i> -acetyl- β -D-glucosaminidase
ER	Endoplasmic reticulum
ERAD	ER-associated degradation pathway
ERV	ER-derived vesicle
Fc	Immunoglobulin ‘Fragment crystalizable’
Fuc	Fucose
FucT	β 1,2-fucosyltransferase
Gal	Galactose
GBP	Green fluorescent protein binding protein
GBP-Fc	GBP coupled at its C-terminus to the murine IgG3 Fc-domain
GFP	Green fluorescent protein
GlcNAc	N-acetylglucosamine
GMI - II	Golgi- α -mannosidase I and II
GnTI - V	N-acetylglucosaminyltransferase I to V
GOI	Gene of interest
GP	Glycoprotein
GP2 – 5	Glycoprotein 2 to 5
HCl	Hydrochloric acid
HRP	Horseradish peroxidase
Hex	<i>Trichoderma reesei</i> β -N-acetylhexosaminidase
HEXO1, HEXO3	β -N-acetylhexosaminidase 1 and 3
IgA, E, G, M	Immunoglobulin isotype A, E, G and M
IMAC	Immobilized metal affinity chromatography
JB man	Jack bean broad spectrum α 1,2,3,6-mannosidase
KDa	Kilodalton
LB	T-DNA left border
LV	Lelystad virus
Lsx	<i>A. thaliana</i> β 1,2-xylosyltransferase <i>medial</i> -Golgi targeting signal
Lsx:endoT	EndoT N-terminally fused to the Lsx

MAb	Monoclonal antibody
MALDI-TOF	Matrix-assisted laser desorption/ionisation–time of flight analyzer
Man	Mannose
NH ₄ HCO ₃	Ammonium hydrogen carbonate
<i>NptII</i>	Neomycin phosphotransferase II gene conferring kanamycin resistance
OD280	Optical density measured at 280 nm
PMP	Plant-made pharmaceutical
Pnos	Promoter of the nopaline synthase gene
Pphas	Promoter of the β-phaseolin gene
PPP	Plant-produced pharmaceutical
Ppt	Phosphinothricin
PPV	Plant-produced vaccine
PRRSV	Porcine reproductive and respiratory virus
PSV	Protein storage vacuole
RB	T-DNA right border
RNA	Deoxyribonucleic acid
ScFv	Single-chain variable fragment
ScFv-Fc	ScFv fused to the Fc-domain of an immunoglobulin
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SS	Signal sequence of the Arabidopsis 2S albumin
T-DNA	Transfer DNA
TFA	Trifluoroacetic acid
Tm	Transmembrane domain
TMB	3, 3', 5, 5'-Tetramethylbenzidine
TMV	Tobacco mosaic virus
TSP	Total soluble protein
Vacc	Vaccinated
WT	Wild type

Xyl

Xylose

Ω

Leader sequence of the 'tobacco mosaic' virus

Scope & objectives

The work presented in this thesis is framed within the larger ambition of the research group headed by Prof. Dr. Depicker to develop a seed-based platform for the production of recombinant proteins. Plants show distinct advantages over other systems for the production of recombinant proteins, which nowadays constitute one of the fastest growing groups of therapeutic products for human use. Plants are capable of complex protein production, lack human and animal pathogens and, for most systems, production is easily scalable and might also be very cheap. Previously, a cassette was developed in our research group for the high, seed-specific expression of a transgene based on regulatory elements that were derived from seed storage proteins of the common bean *Phaseolus vulgaris*. This cassette is used to transform the model organism *Arabidopsis thaliana*, whose small stature facilitates its cultivation in a greenhouse, and promote high seed-specific accumulation of heterologous proteins. The combination of both of these elements, i.e. the cassette and the *Arabidopsis* host, constitutes the seed-platform that is our system of choice and has successfully been used to express a multitude of recombinant proteins ranging from enzymes over single-chain variable fragments to entire antibodies. During this PhD work, two research projects were undertaken in parallel to amplify the capacity of the *Arabidopsis* seed-platform and these are detailed in the following paragraphs.

A first project was aimed at extending the range of proteins produced in our system towards viral antigens, which are complex proteins that are often difficult to express. If their efficacious production could be managed in plants, this would allow plant-based platforms to consolidate a unique position among the conventional, cell-culture based systems. Several proteins of the porcine reproductive and respiratory virus, currently the most important pathogen of swine and for which novel vaccines are highly needed, were selected, extensively optimized and expressed in our system with the intention of evaluating the capacity of the seed-platform for subunit vaccine production.

Despite all the advantages offered by plants for the production of heterologous proteins, they suffer from one drawback, namely that their N-glycosylation machinery differs from that of humans so that therapeutics produced in plants often carry immunogenic glycans, which is potentially very dangerous. This has sparked the development of a plethora of strategies to ‘humanize’ plant glycans and avoid immunogenic glycan residues. An additional problem associated with complex glycans in the context of therapeutic protein production is that these are heterogeneous and lead to the generation of different ‘glycoforms’ of the same protein that show different characteristics, which is highly undesirable from both a regulatory and commercial point of view. To avoid glycan

heterogeneity and immunogenicity, a straightforward method was conceived during the second project, circumventing extensive modifications of the plant glycosylation pathway, to glycoengineer the existing seed-platform so that therapeutic glycoproteins could be produced at high homogeneity while lacking immunogenic plant glycans. The objectives of each project are summarized below.

Vaccine production in *Arabidopsis* seeds:

- Optimization and expression of viral antigens in the platform
- Assessment of the protective capacity of the seed-produced antigens via immunization trials

Glycoengineering in *Arabidopsis* seeds:

- Adapt the platform towards the production of glycoproteins lacking immunogenic plant glycans
- Deliver proof-of-concept by expression of a reporter protein
- Deliver proof-of-functionality by assessment of the immunogenicity of the reporter protein

Summary

The discovery of vaccination, or the deliberate exposure to a pathogen or part thereof with the intention of creating long lasting protective immunity, has given mankind a powerful tool to combat infectious diseases. The advent of vaccination, together with widespread access to clean water, has drastically reduced the incidence of these diseases during the last century. Nowadays, recombinant technologies based on genetic engineering are readily at hand that allow the production of vaccines in an array of cell-culture systems based on bacteria, yeasts, insect or mammalian cells. Alternatively, also plants can be employed to make foreign proteins (on which vaccines are based), a concept called 'molecular farming', and they have a unique mix of features that is useful for the latter. Plants are more evolved than bacteria and yeasts and as such are capable of producing complex proteins very similar (but not identical) to those of humans, they are considered safe because they do not harbor mammalian pathogens, production can easily be amplified by increasing the acreage of plants and no high-tech equipment is needed for their cultivation so protein production is potentially very cheap.

In the research group of Prof. Dr. Depicker, a system was developed to stimulate the high accumulation of heterologous proteins (proteins that are foreign to the host in which they are introduced through genetic engineering) for various uses in the seeds of the model plant *Arabidopsis thaliana*, which is well suited for both lab and commercial scale production of these plant-made pharmaceuticals (PMPs). This system consists of gene elements of storage proteins that are present in very high quantities in the seeds of the common bean *Phaseolus vulgaris*, which are used to drive the expression of a foreign gene upon its introduction in *Arabidopsis*, resulting in high accumulation of the corresponding protein in the seeds. An advantage about the production of PMPs in seeds is that this does not interfere with the growth phase of the plant, the PMP is contained in a compact, desiccated volume and the seeds can be stored for later processing of the PMP. Throughout the years several different proteins have been successfully generated using this system such as enzymes and antibodies and, in collaboration with the research group of Prof. Dr. Nauwynck at the Faculty of Veterinary Medicine, we wanted to employ the benefits of the system for the production of a viral subunit vaccine (a vaccine consisting of specific proteins of a pathogen called antigens). Viral antigens often are complex proteins, which can be difficult to produce and plants might just be very suited for this. The porcine reproductive and respiratory virus (PRRSV) was selected as a target because it is currently the most important pathogen of swine and there is a continuous demand for new vaccines to combat it.

PRRSV antigens had been made in plants before and turned out to be difficult to produce, only accumulating to very low levels in the plant tissues, so we selected glycoprotein 3, 4 and 5 (GP3, 4

and 5) and extensively optimized these with the intention of increasing their accumulation in the seeds of *Arabidopsis*. We removed their transmembrane domains that complicate production and coupled the antigens to stabilizing protein domains such as the green fluorescent protein (GFP) and immunoglobulin G 'crystallizable fragments' (Fc-domains). The resulting set of antigenic proteins, 17 in total, was produced in the seeds and several antigens accumulated to high levels. These could be extracted and purified and when injected in mice, they generated antibodies, some of which had virus-neutralizing capacity, which is considered to be an important parameter for protective immunity against the PRRSV. However, these results could not be repeated during a pilot experiment where three piglets received an antigen cocktail. The exact reason for this is currently not clear and further investigation is required to clarify if the seed-produced antigens can protect piglets *in vivo* from the PRRSV. All in all, we showed expression of the PRRSV antigens to levels far above those achieved by other groups and this demonstrates the suitability, at least concerning accumulation and purification, of our seed-system for the production of viral antigens. However, to assemble an efficacious vaccine against the PRRSV, the seed-produced antigens also need to induce protective immunity and their capacity to do so, remains unclear.

As mentioned before, plants show distinct advantages over other systems for the production of recombinant proteins. The technology has been adapted by major pharmaceutical companies and several products are in clinical trials on the way to be marketed. However, one major drawback is that plant and human N-glycans (sugar structures attached to proteins) differ and although the plant N-glycosylation pathway is to a large extent similar to that of humans, both pathways diverge during the later stages where complex glycans are formed. Complex plant glycans can bear groups that are absent (or present in other linkages) on human glycans and these are controversial because of their immunogenic potential (propensity to provoke a reaction from the immune system). A large fraction of the human population has antibodies in their blood directed against these groups. Although the implications of this observation are not entirely clear and are currently under debate, it is generally accepted that topical application of a PMP decorated with complex plant glycans is safe. However, this picture changes when a therapeutic protein intended for injection in the blood stream is envisaged. Here, binding of these antibodies to the glycans present on the protein might accelerate the clearance of the protein from the blood, but in the worst case scenario it might also cause a severe allergic reaction. Because of these potentially serious consequences a lot of research is dedicated to 'humanizing' plant glycans to avoid these immunogenic glycan groups.

Another complication related to complex glycans in the context of therapeutic protein production is that these constitute a diverse group of structures. A glycoprotein decorated with these structures will thus exist as several 'glycoforms', having the same amino acid backbone but decorated with

different glycan structures. This heterogeneity is highly undesirable from both a commercial and regulatory point of view. To tackle these issues, i.e. glycan heterogeneity and immunogenicity, we devised a simple strategy in collaboration with the research group of Prof. Dr. Callewaert at the Faculty of Sciences to engineer our system so that glycoproteins (proteins decorated with glycans, which are sugar structures) could be produced at great uniformity without immunogenic plant glycans and while maintaining their native fold (normal structure). To achieve this, we introduced a fungal enzyme, the endoT, into the seeds of an *Arabidopsis* mutant, called the *complex-glycan-deficient (cgl)* mutant. This mutant only makes 'high-mannose' glycans and although these are non immunogenic by nature, they are still heterogeneous and additionally, they can target a protein to mannose receptor bearing cells which, except for specific applications, is detrimental. However, these high-mannose glycans are also an excellent substrate for the endoT that cleaves these glycans from proteins leaving only a minimal glycan structure (a single N-acetylglucosamine (GlcNAc) residue) attached to the protein. Because this single GlcNAc is in direct contact with the protein it is important to maintain the overall structure of the protein that, at the same time, will display a completely homogeneous glycan profile devoid of immunogenic plant glycans.

We produced the endoT in the seeds of the *cgl* mutant where it is enzymatically active and introduced the activation-associated secretory protein 1 (ASP1), an antigenic protein from the parasitic nematode *Ostertagia ostertagi*, as a reporter protein. The ASP1 was purified from both wild type (unmodified) *Arabidopsis* seeds and endoT containing seeds, called wild type and GlycoDelete ASP1 respectively, and extensively characterized. This revealed that when the ASP1 is produced in our unmodified system it is indeed decorated with plant specific glycans that are potentially immunogenic, whereas in the endoT seeds these glycans are completely trimmed down to single GlcNAcs. When rabbits were immunized with both ASP1 glycosylation variants, most, but not all rabbits that had received the wild type ASP1 generated antibodies against the plant glycans thus confirming their immunogenic potential. On the contrary, no such antibodies were present among the rabbits that had received the GlycoDelete ASP1, thereby proving that our engineered system can effectively be used to produce glycosylation variants of therapeutic proteins lacking immunogenic glycans. The ASP1 apparently accumulated less in the endoT seeds than in the wild type seeds which could be due to the presence of the endoT gene, the enzyme itself or be related to the integration of the ASP1 gene in the *Arabidopsis* genome. This could, however, also be due to mere clonal variation and further investigation is required to shed light on this matter.

In summary, we produced enzymatically active endoT in seeds of the *cgl* mutant. Upon introduction of the antigenic reporter protein ASP1 in the endoT seeds, its glycan were completely trimmed down to single GlcNAc residues that are in close contact with the protein. When rabbits were immunized

with purified ASP1, we showed that the ASP1 from endoT seeds did not provoke an antibody response against plant glycans, whereas the ASP1 from unmodified, wild type seeds did, thereby proving that the engineered seeds can effectively be used as a platform to produce PMPs with minimal glycans that lack immunogenic groups, which has great economic potential.

Samenvatting

Met de ontdekking van vaccinatie, ofwel de bewuste blootstelling aan een pathogeen (of gedeelte) daarvan met de intentie om langdurige beschermende immuniteit op te wekken tegen dat pathogeen, heeft de mensheid een zeer doeltreffend middel in handen gekregen om besmettelijke ziekteverwekkers te bestrijden. Gedurende de twintigste eeuw heeft vaccinatie, samen met wijdverspreide toegang tot proper water, het voorkomen van deze ziektekiemen drastisch verlaagd. Tegenwoordig laten allerhande recombinante technologieën gebaseerd op genetische modificatie toe om vaccins aan te maken via verschillende systemen die beroep doen op bacteriële-, gist-, insect- en zoogdiercellen. Planten vormen een alternatief voor deze klassieke productieplatformen en zijn ook in staat eiwitten (waarop vaccins gebaseerd zijn) aan te maken, een concept dat men 'molecular farming' noemt. Daarenboven vertonen planten een unieke combinatie van eigenschappen die voordelig is voor de productie van eiwitten: ze zijn namelijk meer geëvolueerd dan bacteriën en gisten wat maakt dat zij in staat zijn om complexe eiwitten aan te maken sterk gelijkend (maar niet identiek, later meer daarover) op hoe deze in het menselijk lichaam aangemaakt worden. Verder is de productie flexibel, ze kan namelijk eenvoudig verhoogd worden door het areaal aan planten uit te breiden en is er voor de cultivatie van planten geen hoog-technologisch materiaal benodigd zodat de productie mogelijks ook zeer goedkoop is.

In de onderzoeksgroep van Prof. Dr. Depicker is een systeem ontwikkeld om heterologe eiwitten (eiwitten die vreemd zijn aan het organisme waar ze via genetische manipulatie in geïntroduceerd worden), voor verscheidene doeleinden, aan te maken in de zaden van het modelorganisme *Arabidopsis thaliana*, dat geschikt is om plant-geproduceerde pharmaceutica (PPPs) aan te maken zowel op laboratorium- als op commerciële schaal. Dit systeem bestaat uit genelementen overgenomen van opslageiwitten die in grote hoeveelheden aanwezig zijn in de zaden van de gewone boon *Phaseolus vulgaris* en die gebruikt worden om de expressie van een vreemd gen te sturen wanneer dit binnengebracht wordt in *Arabidopsis*, zodanig dat het heterologe eiwit sterk accumuleert in de zaden. Voordelig aan de zaadspecifieke aanmaak van een PPP is dat dit de vegetatieve fase (groeifase) van de plant niet beïnvloed, daarenboven accumuleert het PPP in een droog en compact volume en kan het zaad opgeslagen worden zodat het PPP later verwerkt kan worden. Doorheen de jaren zijn verscheidene eiwitten zoals enzymen en antilichamen met succes via dit systeem tot expressie gebracht en wij wilden, in samenwerking met de onderzoeksgroep van Prof. Dr. Nauwynck aan de Faculteit Diergeneeskunde, de voordelen die het zaadplatform ons biedt aanwenden voor de productie van een vaccin. Meer bepaald een vaccin gebaseerd op virale eiwitten (virale antigenen). Virale eiwitten zijn complexe, vaak moeilijk aan te maken eiwitten en planten

zouden wel eens uitermate geschikt kunnen zijn voor hun productie. Het porcien reproductief en respiratoir syndroom virus (PRRSV) werd geselecteerd als doelorganisme, omdat dit momenteel de belangrijkste varkenspathogeen is en er een grote nood is aan nieuwe vaccins om het virus te bestrijden.

PRRSV antigenen werden reeds aangemaakt in verschillende planten met slechts lage antigenopbrengsten als resultaat. De glycoproteinen (eiwitten waar suikerstructuren aan gekoppeld zijn) 3, 4 en 5 (GP3, 4, 5) werden door ons geselecteerd en uitgebreid geoptimaliseerd met als doel om hun accumulatie in de zaden te verhogen. De transmembraan domeinen van de antigenen, die productie bemoeilijken, werden verwijderd en de antigenen werden gekoppeld aan stabiele eiwitdomeinen zoals het 'groen fluorescerende eiwit' en de 'kristalliseerbare fragmenten' van immunoglobulines. De set antigenische eiwitten die zo gevormd werd, 17 in totaal, werd aangemaakt in de zaden en verscheidene antigenen vertoonden een hoge opbrengst. Deze konden geëxtraheerd en opgezuiverd worden en na injectie in muizen waren deze in staat antilichamen op te wekken met de capaciteit om het virus te neutraliseren, wat verondersteld wordt een belangrijke parameter te zijn voor beschermende immuniteit tegen het PRRSV. Deze resultaten konden echter niet herhaald worden tijdens een pilootexperiment waarbij drie biggen een antigencocktail toegediend kregen. De precieze reden hiervoor blijft tot nu toe onduidelijk en verder onderzoek is nodig om uit te maken of de antigenen aangemaakt in *Arabidopsis* zaad in staat zijn om biggen effectief bescherming te bieden tegen het PRRSV. Alles bij elkaar genomen werd de productie van PRRSV antigenen in *Arabidopsis* zaden aangetoond en werden opbrengsten bereikt die vele malen hoger liggen dan voordien ooit gerapporteerd werd, wat bewijst dat het zaadplatform geschikt is, wat betreft accumulatie en zuivering, voor de aanmaak van virale antigenen. Vooralsnog blijft het echter onduidelijk of de antigenen protectieve eigenschappen bezitten en effectief gebruikt kunnen worden om een doeltreffend vaccin tegen het PRRSV samen te stellen.

Zoals eerder vermeld beschikken planten over specifieke voordelen ten opzichte van andere systemen wat betreft de aanmaak van recombinante eiwitten. Deze technologie werd ondertussen opgepikt door voornamelijk farmaceutische bedrijven en verscheidene PPPs bevinden zich in klinische proeven op weg naar commercialisatie. Eén nadeel is echter dat plant en menselijke N-glycanen (suikerstructuren vastgehecht aan eiwitten) verschillen en hoewel de plant N-glycosylatie machinerie sterk lijkt op de menselijke, verschillen beide processen in de afwerking van de glycanen, zodanig dat complexe (afgewerkte) plant suikers structuren dragen die niet of anders voorkomen op menselijke glycanen. Deze plant-specifieke structuren zijn controversieel omwille van hun potentiële immunogeniciteit (hun intrinsieke vermogen om een reactie van het immuunsysteem uit te lokken). Inderdaad, een groot deel van de menselijke populatie draagt antilichamen die tegen deze structuren

zijn gericht. Alhoewel de implicaties van deze observatie momenteel nog onduidelijk zijn en een bron van discussie vormen, wordt algemeen aanvaard dat oppervlakkige behandeling met een PPP gedecoreerd met plant glycanen veilig is. Dit beeld verandert echter totaal wanneer een PPP bedoeld voor injectie in de bloedstroom beschouwd wordt. In dit geval zal de binding van antilichamen gericht tegen plant glycanen namelijk kunnen leiden tot een versnelde verwijdering van het PPP uit de bloedsomloop en in het slechtste geval, mogelijks een zeer ernstige allergische reactie teweeg brengen. Omwille van deze potentiële levensbedreigende gevolgen wordt er dan ook veel onderzoek verricht naar het 'vermenselijken' van plant glycanen met de intentie de immunogene structuren te vermijden.

Een bijkomend probleem in de context van therapeutische proteïnen is dat complexe glycanen een verzameling van heterogene structuren vormen. Dit zorgt ervoor dat een eiwit zal voorkomen als een set van 'glycovormen' (eenzelfde eiwit maar gedecoreerd met verschillende glycaanstructuren) die verschillende eigenschappen zullen vertonen en dus elk apart gekarakteriseerd moeten worden, hetgeen de commercialisatie en regularisatie van een therapeutisch glycoproteïne verder bemoeilijkt. Om een mouw te passen aan deze heterogeniteit en de eerder vermelde immunogeniciteit, werd in samenwerking met de onderzoeksgroep van Prof. Dr. Callewaert aan de Faculteit Wetenschappen een eenvoudige strategie uitgedacht om het zaadplatform aan te passen, zodanig dat glycoproteïnen aangemaakt kunnen worden met een homogeen glycaanprofiel zonder immunogene plant glycanen en met behoudt van hun originele structuur. Om dit te bereiken werd een enzyme van de schimmel *Trichoderma reesei*, het endoT, geïntroduceerd in de zaden van de *Arabidopsis complex-glycan-deficient (cgl)* mutant die enkel 'hoog-mannose' glycanen aanmaakt. Hoewel deze glycanen van nature uit niet immunogeen zijn, is ook deze groep heterogeen. Daarenboven kunnen hoog-mannose glycanen door hun interactie met de mannose receptor een eiwit naar een specifieke set cellen sturen wat, op enkele gerichte applicaties na, liefst vermeden wordt. Deze hoog-mannose structuren zijn dus niet wenselijk. Het endoT is echter in staat deze glycanen af te splitsen van eiwitten waarbij slechts een minimale glycaanstructuur op het eiwit achterblijft die bestaat uit 1 enkele N-acetylglucosamine suikergroep (GlcNAc). Aangezien deze in direct contact staat met het eiwit is deze belangrijk om de structuur van het eiwit te behouden, maar door de glycaanstructuur te reduceren tot enkel en alleen deze GlcNAc groep wordt tegelijkertijd ook een compleet homogeen glycaanprofiel verkregen dat daarenboven ook nog eens geen immunogene plant glycanen draagt.

Het endoT werd aangemaakt in de zaden van de *cgl* mutant waar werd aangetoond dat het enzyme actief is. Vervolgens werd het 'activatie-geassocieerd secretorisch eiwit 1' (ASP1), een antigeen eiwit van de parasitaire nematode *Ostertagia ostertagi*, ingebracht in de endoT zaden als reportereiwit.

Het ASP1 werd opgezuiverd uit zowel niet-gemodificeerd wild type *Arabidopsis* zaad als uit het endoT zaad, wild type en GlycoDelete ASP1 respectievelijk, en extensief gekarakteriseerd, wat aan het licht bracht dat het wild type ASP1 inderdaad potentieel immunogene plant glycanen draagt, daar waar de glycanen aanwezig op het GlycoDelete ASP1 volledig gereduceerd waren tot GlcNAc residus. Beide ASP1 varianten werden gebruikt om konijnen te immunizeren waarbij de meerderheid, maar niet alle konijnen die geïnjecteerd waren met het wild type ASP1 antilichamen aanmaakten specifiek gericht tegen plant glycanen, waardoor de immunogeniciteit van de plant glycanen nogmaals bevestigd werd. Deze antilichamen konden echter niet gedetecteerd worden in de sera van de konijnen die het GlycoDelete ASP1 hadden toegediend gekregen, waarbij we aantoonen dat het aangepaste zaadplatform effectief gebruikt kan worden voor de aanmaak van PPPs zonder immunogene plant glycanen. De opbrengst van het ASP1 eiwit in de endoT zaden bleek lager te liggen dan dit bereikt in de wild type zaden en dit is mogelijks toe te schrijven aan aanwezigheid van het endoT gen, het endoT enzyme zelf of dit is een artefact te wijten aan de integratie van het ASP1 gen in het genoom van de *Arabidopsis cgl* mutant. Het zou echter ook kunnen dat deze observatie aan louter toeval toe te schrijven is en verder onderzoek is nodig om deze kwestie verder uit te spitten.

Om samen te vatten kunnen we stellen dat actief endoT werd aangemaakt in de zaden van de *cgl* mutant en dat dit in staat was om de glycanen van het antigenische reportereiwit ASP1 dat nadien geïntroduceerd werd in de endoT zaden volledig te trimmen tot enkel de GlcNAc residus die in direct contact staan met het eiwit. Toen konijnen geïnjecteerd werden met opgezuiverd ASP1 geproduceerd in de endoT zaden, induceerden deze structuren geen antilichamen specifiek voor plant glycanen, daar waar de glycanen aanwezig op the ASP1 afkomstig uit de wild type zaden dat wel deden, waarbij bewezen werd dat de endoT zaden weldegelijk gebruikt kunnen worden als productieplatform voor de aanmaak van PPPs met minimale, niet-immunogene glycanen, wat een groot commercieel potentieel heeft.

Chapter I

**Towards a viral subunit vaccine produced in *A. thaliana* seeds,
an introduction.**

Robin Piron and Ann Depicker

R.P. wrote the chapter and A.D. edited it.

Vaccination

The early days

Human history has been entangled with that of infectious diseases, outbreaks of which would periodically decimate or even topple entire societies. It is estimated that during the 14th century the plague, also known as “Black Death”, killed one out of three Europeans (Perry & Fetherston, 1997). Two centuries later the arrival of Europeans in Central America, carrying with them diseases that had been raging in Europe but were foreign to the Americas, initiated the demise of many great Mesoamerican civilizations. Disease spread rapidly, killing thousands, crippling societies and paving the way for the conquistadores that followed in its footsteps (Diamond, 1997). Mankind was left at the mercy of these infectious diseases and their devastating consequences until vaccination, the deliberate exposure to a pathogen or part of it to immunize against that pathogen, made its entry in human history. It was Jenner’s observation that milkmaids exposed to cowpox were protected from smallpox that sparked the first large scale vaccination program in 1798 (Bazin, 2011). This eventually culminated in the eradication of the smallpox in 1979 (Fenner et al., 1988). At the end of the 19th century the pioneer Louis Pasteur discovered that microorganism could be attenuated by exposure to environmental stresses (Plotkin, 2005). This initiated a new era during which vaccinology, the science behind vaccination, rapidly expanded. Pasteur’s principles of ‘isolate, inactivate and inject’ served as key concepts during this era and the initial discovery of attenuation was rapidly followed by the development of methods to inactivate whole microbes and the discovery of bacterial toxins and how these could be manipulated to produce toxoids, their harmless but immunogenic counterparts (Plotkin & Plotkin, 2011; Rappuoli, 2007; Rinaudo et al., 2009). The importance of the work of these early scientists is reflected in the fact that the main vaccines in use nowadays are still based on the three concepts of attenuation, inactivation and administration of components (subunits) of pathogens to induce immunity.

Attenuated, inactivated and subunit vaccines

The purpose of vaccination is to induce long-lasting protective immunity from the exposure to pathogens or parts thereof, with the intention of preventing clinically relevant infectious diseases (Lambrecht et al., 2009). Until today, three major classes of vaccines have been in use to fight these infectious diseases, i.e. attenuated, killed and subunit vaccines (Buonaguro & Butler-Ransohoff, 2010), each with its own merits and drawbacks. Attenuated vaccines consist of live pathogens (or infectious virions in the case of viruses) that have been weakened or adapted to growth in a non-host

organism prior to their inclusion in the vaccine (Plotkin, 2005). As a consequence of this, the pathogen will only be capable of suboptimal replication after injection and will cause a mild infection instead of the full-blown disease. As soon as an adaptive immune response is mounted, the pathogen is effectively eliminated from the host. The biggest advantage of this type of vaccines is that they tend to induce long lasting immunological memory, with the obvious downside being that the pathogen effectively replicates in the host organism and on occasion, this can cause serious disease. This is why, whenever feasible, inactivated vaccines or subunit vaccines, are preferred over attenuated vaccines. Neither inactivated vaccines, nor subunit vaccines contain infectious pathogen and hence are safe, but they both have their own drawbacks. The procedures used to inactivate pathogens can alter the structural elements that are important to mount an effective immune response. Additionally, the induced immune response will be weaker because of lack of replicating pathogen and the combination of these two factors causes inactivated vaccines to be less potent than their attenuated counterparts. Subunit vaccines on the other hand, contain only those elements of a pathogen that are known to be important to obtain protective immunity (Plotkin, 2005). However, without replicating pathogen or the co-stimulatory molecules that are still present in killed pathogens, the antigens used in subunit vaccines are generally very weak immunogens and need to be administered with strong immune stimulating substances called adjuvants (Granell et al., 2010; Lambrecht et al., 2009). Which type of vaccine is used against a specific pathogens depends on the intrinsic characteristics of the pathogen but, given that they are sufficiently potent to induce lasting immunological memory, the tendency is towards inactivated and subunit vaccines, with the latter being favored by the technologies nowadays at hand for recombinant protein production.

Recombinant vaccines

The advent of molecular biology and genetic engineering in the 1980s revolutionized the production of vaccines and especially that of subunit vaccines. Whereas previously tedious procedures had to be followed to isolate specific components from pathogens, obviously with the risk of contamination by live pathogen, genetic engineering now allowed the production of selected antigenic proteins in bacteria, yeasts, animal and insect cells (Plotkin & Plotkin, 2011). One of the first successes of this new technology was a hepatitis B vaccine based on a recombinant hepatitis B surface antigen produced in yeast which substituted purification of the surface antigen from plasma of infected individuals (McAleer et al., 1984). Another famous example attained through genetic engineering are the human papilloma virus vaccines to protect against cervical cancer (Frazer, 2004). Here the major capsid protein of the virus auto-assembles into highly immunogenic virus like particles when it is

produced in eukaryotic cells (Kirnbauer et al., 1992). By now numerous vaccines against a whole range of pathogens have been developed and their distribution, together with widespread access to purified water, have reduced the incidence of infectious diseases causing mortality among children by 97 to 99% (Rappuoli et al., 2002). The world health organization estimates that vaccines avert two to three million deaths a year making it the single most cost-effective tool to combat infectious diseases. In this respect vaccination should be considered no less than one of the major achievements of mankind.

Transgenic plants as biofactories

The demonstration in 1983 by Herrera-estrella et al. (1983) that a foreign resistance gene could be inserted into tobacco plants by agrobacterium-mediated transformation and that these subsequently made functional protein, opened up a whole new field for the production of therapeutic proteins. Hopes were high as plants have some distinct features as compared to the more conventional systems based on cultured bacteria, yeast or mammalian cells. Plants were considered to be safe as they are not susceptible to human or animal pathogens and to be capable of complex protein production, given that they correctly perform post-translational modification such as disulfide bond formation and N-glycosylation (Gomord & Faye, 2004), a presumption that was soon confirmed when Hiatt et al. (1989) showed the correct production of a recombinant antibody in tobacco plants. Additionally it was expected that plants represented a cheaper system, given that less high-tech equipment would be needed for their cultivation. Rapidly a future was envisaged in which plants would both feed and heal mankind. Edible vaccines such as bananas or tomatoes with curative properties would be grown cheaply where they were needed most, obviously with great benefit to the poor and sick in the world (Ma et al., 2003; Penney et al., 2011). The field of 'molecular farming', or the production of plant-made-pharmaceuticals (PMPs), expanded quickly and many of the proteins previously produced in conventional cell culture systems were introduced into plants. Antigens from major pathogens such as rabies, *Yersinia pestis*, the hepatitis B virus, human papilloma virus and human immunodeficiency virus (HIV) were all produced in plants and complex structures such as bacterial toxins, virus-like particles (VLPs) and antibodies were shown to assemble correctly (Alvarez & Cardineau, 2010; Giorgi et al., 2010; Shchelkunov & Shchelkunova, 2010). Proof-of-concept of the immunogenicity of several plant-produced antigens was also delivered in animal models and when Thanavala et al. (2005) demonstrated that human immunity against hepatitis B virus could be boosted by feeding transgenic potatoes, the dream of edible vaccines for the poor became almost tangible. However, this initial enthusiasm was tempered by practical constraints and the same study

also highlights these shortcomings. Yields of antigen were low, a recurring issue for many PMPs (Nagels et al., 2012b; Pelosi et al., 2012), requiring the human volunteers to consume quite large volumes of potato pointing towards a second drawback, namely that of dose-standardization (Paul & Ma, 2010). To receive the correct vaccine dose patients should ingest a defined quantity of transgenic material, a prerequisite that would be difficult to implement outside of a laboratory setting. This eventually led to the consensus that plant-produced-vaccines (PPVs) need to be processed into a convenient formulation containing a standardized dose so they can be easily administered (Rybicki, 2009).

Currently, different types of vaccines against a range of pathogens are being investigated and developed. Oral (mucosal) vaccines are still a major research focus and formulations have been devised with excellent thermo stable properties that allow the storage of the vaccine at ambient temperature for prolonged periods of time (Penney et al., 2011). A property that has led to the replacement of the term 'edible vaccine' with the more appropriate 'heat stable oral vaccine' (Rybicki, 2010). Another major research focus is on purified vaccines for parenteral delivery since injection remains the most effective route of vaccination to date (Pelosi et al., 2012). Full purification of a PPV provides the greatest control over the administered product and the immune response that is induced is generally stronger so less antigen is needed (Paul & Ma, 2010; Salyaev et al., 2010). The decades of research carried out ever since the conception of PMPs also tackled the initially recurring issue of low protein yield. Although the optimal conditions for a given PMP still need to be identified on a case-by-case basis, different transformation systems are now available for a whole range of plant species and a variety of regulatory elements allow for tissue specific and subcellular targeted accumulation of the PMP. This has led to reported record yields of up to 80% of total soluble protein (TSP) for a transient viral based system (Marillonnet et al., 2004), 51% of TSP after stable chloroplast transformation (Lentz et al., 2010) and 37% of TSP for a seed-produced PMP obtained via stable genomic transformation (De Jaeger et al., 2002).

In 2006, the first PPV was allowed onto the market. The United States Department of Agriculture granted market approval to Dow Agro Science for their veterinary vaccine against the Newcastle disease affecting poultry¹. Dow Agro Science never commercialized the vaccine but used it as proof-of-concept for the suitability of their tobacco suspension cell system for vaccine production. Together with several dedicated plant biotech companies most big pharmaceutical companies now also have a division working on PMPs as exemplified by the acquirement of ICON genetics by Bayer.

¹ http://www.aphis.usda.gov/animal_health/vet_biologics/publications/notice_06_10.pdf

Recently, the Israeli company Protalix was allowed to market its PMP Eleyso² under an expanded orphan disease program. Eleyso is the brand name of the human enzyme glucocerebrosidase produced in carrot cell culture and it is used in enzyme replacement therapy to treat Gaucher's disease, a genetic lysosomal storage disease (Aviezer et al., 2009). With this milestone approval of a non-topical PMP for parenteral use in humans and several PPVs in different stages of clinical trials, both for oral as parenteral delivery, it is expected that PPVs are on the verge of breaking through as valuable and useful alternatives for the current expensive and limited production of vaccines via the classical bacteria, yeast and mammalian cell based culture systems (Penney et al., 2011).

² <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm302549.htm>

Seed production of PMPs

Over the years, many different platforms have been used for recombinant protein production. These include leafy crops, cereal and legume seeds, oilseeds, fruits and vegetables, vegetative tissue of higher plants, cell cultures, algae, mosses and aquatic plants (Ma et al., 2003; Obembe et al., 2011). All these systems have their advantages and disadvantages, but seeds are of particular interest because they have evolved to stockpile proteins for extended periods of time in a stable environment (Stoger et al., 2005). Indeed, when recombinant antibodies and formats derived thereof were produced in seeds they accumulated to high levels, were stable at room temperature for extended periods and retained their binding capacity (reviewed in Viridi and Depicker (2013)). Several factors drive these desirable traits. The abundance of chaperones and disulfide isomerases during seed formation ensures that proteins reach their correct conformation, proteases are little abundant and the compacted, dessicated nature of seeds stabilizes proteins (Muntz, 1998). Seed water content is generally below 10% which is in stark contrast with leaves that in most cases contain more than 90% of water (Boothe et al., 2010). The long term, stable storage of proteins in a small volume eliminates the need for immediate processing of the harvested material. This decoupling of the cultivation phase and further downstream processing gives seeds a marked advantage over other more watery tissues, that need to be processed by freeze-drying upon harvesting or from which the PMP has to be purified immediately. Downstream processing is further facilitated by the fact that the seed proteome is relatively simple and that secondary metabolites known to interfere with purification, such as alkaloids, are only present in small amounts (De Jaeger et al., 2002; Stoger et al., 2005). Finally, seed-specific production of a PMP does not intervene with the vegetative phase of plant growth.

A range of seeds are being employed for commercial PMP production, each with its own characteristics. Maize has the highest seed biomass but has the undesirable trait of outcrossing, increasing the risk for transgene spread. Barley and rice are both self-pollinating and have relative high seed yields. Oilseeds such as safflower and rapeseed allow the PMP to be specifically purified by targeting to the oil bodies that are then easily separated from the crude seed extract and the PMP released, a technique used by the former Canadian company SemBioSys Genetic Inc. Although not being used on a commercial scale, the legumes pea and soybean also have a high potential for PMP production because of their massive protein content that can amount up to 40% of seed weight (Boothe et al., 2010; Stoger et al., 2005).

Another interesting species, *Arabidopsis thaliana*, has some distinct features beneficial for PMP production that are not encountered with the seed crops mentioned before. As a model organism

protocols for transformation and optimal growth are widely available. It is self-pollinating and its small stature facilitates growing in greenhouses, thus reducing biosafety issues. It has a short life cycle of three months that considerably reduces the timeline between transformation and PMP. Methods for automatic sowing and harvesting of the seeds are under development and until today, some the highest seed yields of PMP have been obtained in *Arabidopsis* (De Jaeger et al., 2002; Loos et al., 2011a).

The *Arabidopsis* seed platform

To take full advantage of the favorable capacities of *Arabidopsis* for the production of PMPs, an expression cassette was developed in our group designed for high seed-specific expression of recombinant proteins. It was hypothesized that the regulatory elements that are responsible for the massive accumulation of the endogenous seed storage proteins of the dicotyledonous plant *Phaseolus vulgaris* would bestow the same characteristic onto a heterologous protein if placed under control of these elements. Initially it was proven that the arcelin seed storage protein from the common bean (*Phaseolus vulgaris*) also accumulated to high levels (15% of TSP) when it was introduced in *Arabidopsis* (Goossens et al., 1999). This spurred the development of an expression cassette based on the promoter and terminator from the arcelin5-I gene and indeed when *Arabidopsis* was transformed with a murine single-chain variable fragment (scFv) cloned into this cassette, the scFv accumulated up to 12.5% of TSP in the seeds of homozygous offspring (De Jaeger et al., 2002). The cassette was further optimized by changing the promoter of the arcelin5-I gene for that of the gene encoding β -phaseolin, another important seed storage protein from *Phaseolus vulgaris*, and substituting the 5'-untranslated region (5'-UTR) of the arcelin5-I gene by that of the tobacco mosaic virus (TMV). Accumulation levels doubled when this optimized cassette was used to drive expression of the same murine scFv and it even established an all-time record yield for a seed-produced PMP by reaching an accumulation level of no less than 36.5% of TSP in the seeds of homozygous transformants (De Jaeger et al., 2002).

The importance of the regulatory elements that make up the expression cassette in achieving the above mentioned yields can hardly be overestimated. Accumulation of scFv-Fcs in *Arabidopsis* seeds under control of the strong, constitutive 35S promoter derived from the cauliflower mosaic virus was considerably less than when the β -phaseolin promoter was used (De Jaeger et al., 2002; Loos et al., 2011b). The scFv-Fcs under control of the latter were also structurally more integer, showing less degradation than their 35S driven counterparts and this has been attributed to the different spatiotemporal expression pattern of both promoters (Loos et al., 2011b). The 35S promoter is active

during the early stages of seed development only whereas the *β -phaseolin* promotor is expressed during the entire period of seed filling and this might expose accumulating recombinant proteins to different sets of proteases, with the expression pattern of the *β -phaseolin* promotor being more beneficial for protein storage and integrity (Chen et al., 1986; Petruccelli et al., 2006; Viridi & Depicker, 2013).

Ever since its conception, the cassette has been used to drive the seed specific expression of a multitude of recombinant proteins in *Arabidopsis* such as scFvs (De Jaeger et al., 2002; Van Droogenbroeck et al., 2007), fusions of scFvs with the “Fragment crystallizable” (Fc domain) of immunoglobulins (IgGs) (Loos et al., 2011b), entire monoclonal antibodies (Loos et al., 2011a), enzymes and auto-antigens (Morandini et al., 2011) and lately also a whole range of nanobodies (VHHs), VHH-Fc fusions (De Buck et al., 2013) and even synthetic, dimeric secretory IgA (Viridi et al., 2013). With some exceptions, the accumulation of the recombinant proteins produced in the seeds of *Arabidopsis* ranged from <0.1 to 15% of TSP. Antibodies and synthetic variants thereof represent the upper end of the range and the *Arabidopsis* seed platform seems to be exceptionally suited for their production, as is exemplified by the record yield achieved by De Jaeger et al. (2002). The lower end of the range is mainly occupied by enzymes and auto-antigens although also VHHs as such do not accumulate very well and this has been attributed some instability at the protein level and/or transcriptional inefficiency (De Buck et al., 2013).

Quite some exceptional yields have been achieved during the years using the seed-specific expression cassette and although it was shown that the accumulation of recombinant antibodies to as little as 1% of TSP is sufficient to induce an unfolded protein response, *Arabidopsis* seems rather tolerant to the high accumulation of heterologous proteins in its seed, showing no dramatic impact on seed germination and seedling growth (De Jaeger et al., 2002; De Wilde et al., 2013). This excellent capacity for the production of recombinant proteins combined with its small stature, short life cycle and its ease of transformation make the *Arabidopsis* seed-platform detailed here our system of choice. Additionally, the seed-platform offers possibilities for the production of recombinant proteins at a commercial scale. Automated sowing and harvesting of seeds is under investigation and it has been estimated that *Arabidopsis* grown in a greenhouse facility would yield between 180 and 240 grams of seeds per square meter on a yearly basis (He et al., 2013; Loos et al., 2011a). Taking 10% of TSP as a realistic value for a product to be commercialized, these yield would translate to (approximately one fifth of the dry seed weight consists of extractable protein) between 3.6 and 4.8 grams of recombinant protein per square meter.

With the excellent track record of the seed-platform as motivation, we wanted to extend the range of recombinant proteins produced towards antigens and evaluate the possibilities the platform offers for the production of subunit vaccines. More specifically, viral subunit vaccines were envisaged because viral proteins are complex and hence, to truthfully reproduce these, an eukaryotic system is needed. Plants could occupy a niche among the current available platforms by producing these exigent proteins. To deliver proof-of-principle of the suitability of the seed-platform for the production of antigenic proteins, the porcine reproductive and respiratory syndrome virus (PRRSV) was chosen as a model organism and a set of its glycoproteins was selected to be expressed in the seed-platform.

The porcine reproductive and respiratory syndrome virus

During the eighties a new disease emerged among swine typified by symptoms related to respiratory distress and reproductive failure. Initially a multitude of names was used to describe the disease such as the 'swine infertility and respiratory syndrome', 'the mystery swine disease' and the Dutch name, 'abortus blauw', based on the observation that infected sows would occasionally present a blue discoloration of the ears (Collins et al., 1992; Wensvoort et al., 1991). Eventually porcine reproductive and respiratory syndrome (PRRS) was adopted as a common name to describe the disease and the causative organism, the porcine reproductive and respiratory syndrome virus (PRRSV) was almost simultaneously isolated by researchers in North America and The Netherlands (Collins et al., 1992; Terpstra et al., 1991; Wensvoort et al., 1991). The virus is classified among the order of the *Nidovirales* in the family of the *Arteriviridae*, together with the simian haemorrhagic fever virus, the lactate hydrogenase elevating virus and the equine arteritis virus, all of which are positive stranded RNA viruses (Meulenbergh et al., 1993b; Snijder & Meulenbergh, 1998). These viruses share many characteristics such as their genome organization, gene expression strategy, and the tendency to induce persistent infections (Meulenbergh et al., 1997b). They also primarily infect cells of the monocyte/macrophage lineage and/or endothelial cells. Since its emergence, the PRRSV has spread around the globe and is now endemic in most pig-rearing countries of the world. It is considered to be one of the most important pig pathogens worldwide and it is estimated that in the US alone economic losses inflicted by the PRRSV amount to over 500 million dollars yearly (Neumann et al., 2005). Two genotypes exist, an American and North European, which in their turn can be divided into several subtypes (An et al., 2010; Karniychuk et al., 2010).

The genome of the PRRSV is about 15 kilobases (kb) long and contains multiple open reading frames (ORFs) that are transcribed as a nested set of subgenomic mRNAs, a typical feature among the

Nidovirales (Figure 1B). The PRRSV mRNAs are coterminal, sharing their 3'-end and also have a common 5'-leader sequence that is derived from the genome 5' end and is fused to the subgenomic mRNAs by a discontinuous transcription mechanism (Meulenberg et al., 1993a). The first two ORFs that comprise about three quarters of the entire genome, ORF 1a and 1ab, encode two polyproteins that undergo auto-proteolytic processing to generate up to fourteen nonstructural proteins and these are involved in genome replication and immunomodulation of the host organism (Fang & Snijder, 2010; Kimman et al., 2009; Sun et al., 2010). The remaining ORFs code for the structural proteins of the virion and several of these are embedded in the lipid envelope surrounding the nucleocapsid (Figure 1A).

The glycoproteins 2, 3 and 4 (GP2, 3 and 4) that result from translation of the ORFs 2a, 3 and 4, are the minor envelope proteins and are present in the lipid membrane as a non-covalently associated heterotrimer (Meulenberg et al., 1995; Meulenberg & PetersendenBesten, 1996). The major envelope proteins GP5 and the M proteins, encoded by ORF5 and 6 respectively, occur as disulfide bridge-linked dimers (Mardassi et al., 1996). Translation of ORF7 generates the N protein that makes up the nucleocapsid. The three main structural proteins, GP5, M and N are indispensable for viral particle formation whereas GP2, GP3 and GP4 are essential only for virus infectivity (Wissink et al., 2005). Two additional genes are present in the PRRSV genome and these are entirely embedded within the existing ORF2 and 5. Translation of the former yields the small, non-glycosylated E protein, postulated to function as an ion channel during release of the genome in the cytosol, whereas the latter generates the ORF5a protein whose function is unknown (Johnson et al., 2011; Lee & Yoo, 2006; Snijder et al., 1999; Wu et al., 2001).

In nature, the PRRSV is primarily transmitted via the respiratory route but virions are also found in the semen of infected boars which makes sexual transmission another important route of infection (Snijder & Meulenberg, 1998). Additionally, the virus can cross the placenta during late gestation and infect unborn piglets thereby inducing reproductive failure (Rowland, 2010). PRRSV has a tropism for specific subsets of macrophages which it infects in the lungs, lymphoid tissue and placenta (Van Breedam et al., 2010a). Attachment to macrophages is initiated by loose binding of the GP5-M dimer to heparan sulfate moieties that are part of the O-linked glycans carried by glycoproteins anchored in the cell membrane (Delputte et al., 2002). The dimer then stably engages the sialoadhesin receptor via sialic acids present on the GP5 glycans and this triggers the clathrin mediated endocytosis of the virion (Delputte & Nauwynck, 2004; Nauwynck et al., 1999; Van Breedam et al., 2010b). Once internalized, a drop in pH is needed to uncoat the virion and deliver the genome to the cytosol so productive infection can occur. This process is crucially mediated by the interaction of the GP2, 3 and 4 heterotrimer with the scavenger receptor CD163 and depends on the activity of the aspartic

protease cathepsin E, and an unidentified trypsin-like serine protease (Das et al., 2010; Misinzo et al., 2008; Van Gorp et al., 2008).

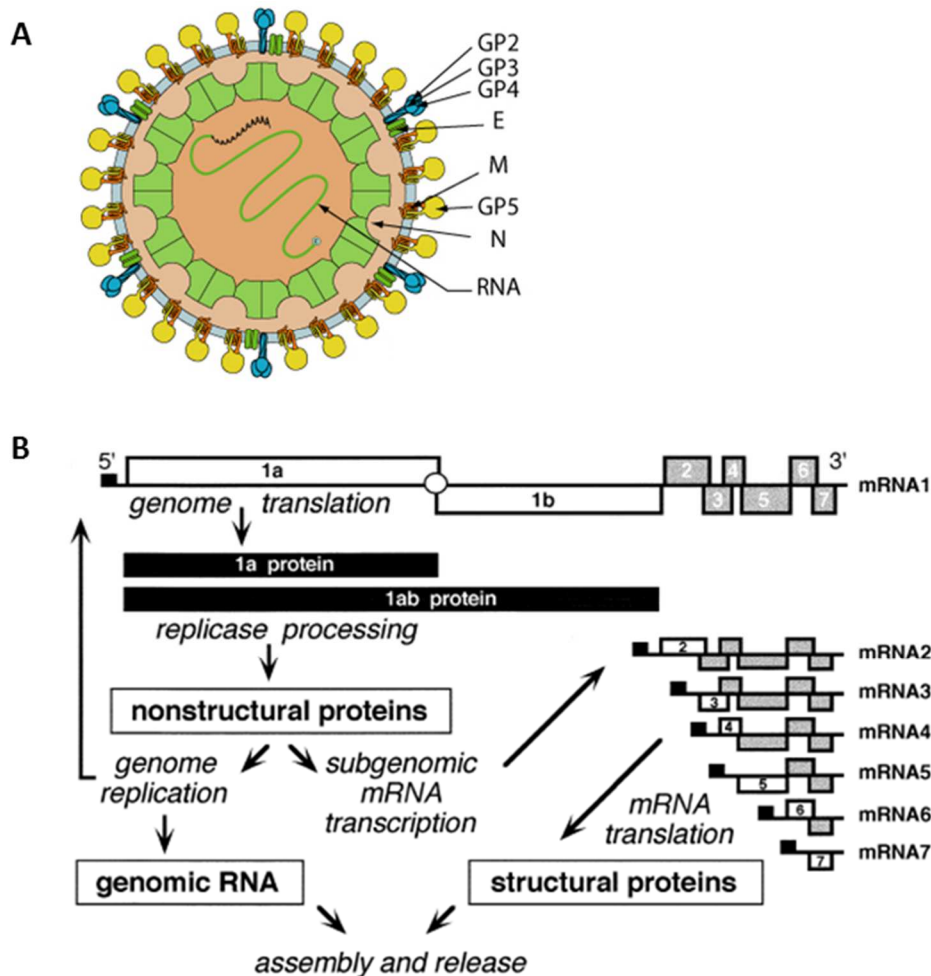


Figure 1. Overview of the PRRSV virion, genome organization and expression. (A) The PRRSV virion³. The single-stranded RNA genome is enclosed in the nucleocapsid consisting of the N protein. The envelope glycoproteins are depicted embedded in the lipid membrane surrounding the nucleocapsid. (B) PRRSV genome organization and expression. The different ORFs and subgenomic mRNAs are depicted together with the polyproteins that give rise to the nonstructural proteins. Genome expression starts with the translation of the 1a and 1ab polyprotein (through a ribosomal frame shift), which are subsequently proteolytically processed into the nonstructural proteins that modulate the host's immunity and are responsible for further viral replication. The different subgenomic mRNA's that generate the structural proteins (the corresponding ORF is given as a white, numbered box) are depicted on the right showing their common 5'-leader sequence (black box). Once all the structural proteins are translated, they associate with the genome into infectious virus particles that are subsequently released from the host cell. Figure adapted from Snijder and Meulenberg (1998).

³ http://education.expasy.org/images/Arteriviridae_virion.jpg

PRRSV causes respiratory distress among pigs of all ages but is especially problematic when infecting pregnant sows leading to abortion, early farrowing and birth of death or weakened piglets (Collins et al., 1992; Terpstra et al., 1991). The situation is further aggravated by the fact that during recent years also more virulent strains have been emerging with a high incidence of morbidity and mortality among pigs of all ages, as is exemplified by the 2006 outbreak in China infecting two million pigs of which 400.000 succumbed (An et al., 2010; Karniychuk et al., 2010; Tian et al., 2007; Zhou & Yang, 2010). When not lethal, infection is very persistent and can take several months before it is completely resolved by the immune system that is apparently not able to cope efficiently with the virus (Figure 2). The initial immune response is crippled and typified by the late appearance of neutralizing antibodies and cell mediated immunity. Several mechanisms have been proposed to explain the delayed immune response mounted against the PRRSV. The virus has immune modulatory properties and interferes with the host's interferon alpha production (Albina et al., 1998; Kimman et al., 2009). Sensitive epitopes on the envelope proteins are protected by glycan shielding (Ansari et al., 2006; Vu et al., 2011). The envelope proteins carry non-neutralizing, immunodominant decoy B-cell epitopes that are targeted by the initial non-effective antibody response and the neutralizing epitopes are often highly variable (Ostrowski et al., 2002; Vanhee et al., 2010).

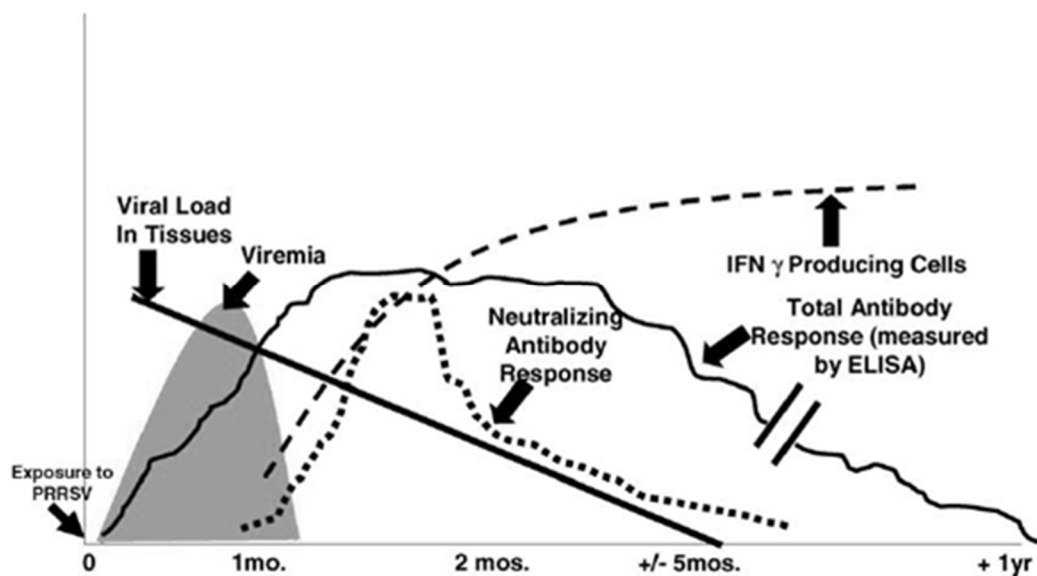


Figure 2. The course of a PRRSV infection. Taken from Lopez and Osorio (2004). PRRSV induces a defective immune response that is initially not able to cope with the virus and is characterized by the delayed onset of an efficient adaptive immune response; it is only after several weeks that neutralizing antibodies are generated and IFN- γ secreting cells arise. The latter two eventually reduce viremia, but the virus nevertheless persists in peripheral lymphoid tissues and it takes months before it is completely cleared from the host. Abbreviations used: mo(s), months; yr, year; IFN- γ , interferon gamma.

To combat PRRSV, prophylactic vaccines are administered that alleviate the clinical symptoms of a subsequent infection. However, their production is hampered by the genetic drift of the virus and it is generally accepted that no single vaccine will ever be capable of entirely controlling the PRRSV (Nauwynck et al., 2012). Strategies are thus being devised that incorporate the use of different types of vaccines to maximally profit from the specific benefit offered by each type. Naïve animals should, for example, be vaccinated with attenuated vaccines that have been shown to provide good protection, whereas inactivated or subunit vaccines should be used to boost the preexisting immunity of sows. The latter are safer and would result in increased levels of neutralizing antibodies in the blood and colostrum that protect sows against the transplacental spread of the PRRSV and the birth of viremic piglets (Nauwynck et al., 2012). With the intention of producing such a subunit vaccine, we selected the GP3, GP4 and GP5 as candidate antigens to be expressed in our seed-platform. Neutralizing antibodies play an important role in clearance of cell-free virus whereas cell-mediated immunity is needed to cope with PRRSV infected cells. The ideal vaccine should thus contain important B and T-cell epitopes. Pepscan analysis of antibodies in the sera of pigs after PRRSV infection revealed that both GP3 and GP4, but not GP5, carry neutralizing B-cell epitopes that were recognized by the sera of multiple pigs and these thus constitute antigens with vaccine potential (Vanhee et al., 2011). Although the GP5 selected in this study does not contain such epitopes, it was shown that it carries multiple important T-cell epitopes and hence, was also retained to be included in the experimental vaccine (Diaz et al., 2009; Vashisht et al., 2008).

Chapter II

**Towards a viral subunit vaccine produced in *A. thaliana* seeds, results:
“Boosting in planta production of antigens derived from the porcine
reproductive and respiratory syndrome virus (PRRSV) and
subsequent evaluation of their immunogenicity”.**

Robin Piron, Stefaan De Koker, Annelies De Paepe, Julie Goossens, Jonah Nolf, Els van Lerberge,
Janne Six, Johan Grooten, Hans Nauwynck and Ann Depicker.

Adapted from:

Boosting In Planta Production of Antigens Derived from the Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) and Subsequent Evaluation of Their Immunogenicity. Piron R, De Koker S, De Paepe A, Goossens J, Grooten J, Nauwynck H, Depicker A (2014) *PLoS ONE* **9**: e91386

R.P. wrote the manuscript and the corresponding chapter. R.P., S.D.K., A.D.P., J.G., H.N. and A.D. conceived the experimental design. R.P., S.D.K., A.D.P., J.G., J.N., E.V.L. and J.S. performed the experiments. A.D.P. generated the GP4, GP4(-Tm), GP5 and GP5(-Tm) constructs and corresponding transgenic lines. E.V.L., J.G. and J.S. helped with the propagation of the transgenic lines and the analysis of the antigen accumulation levels. J.N. assisted during the purification of the antigens. S.D.K. carried out and analyzed the mouse immunization trial. The research group of professor Nauwynck performed the piglet immunization trial and all the IPMA and serum-neutralisation assays. S.D.K., A.D.P., J.G., H.N. and A.D. critically revised the manuscript and approved the final version.

Abstract

Porcine reproductive and respiratory syndrome (PRRS) is a disease of swine, caused by an arterivirus, the PRRS virus (PRRSV). This virus infects pigs worldwide and causes huge economic losses. Due to genetic drift, current vaccines are losing their power. Adaptable vaccines could provide a solution to this problem. This study aims at producing *in planta* a set of antigens derived from the PRRSV glycoproteins (GPs) to be included in a subunit vaccine. We selected the GP3, GP4 and GP5 and optimized these for production in an *Arabidopsis* seed platform by removing transmembrane domains (Tm) and/or adding stabilizing protein domains, such as the green fluorescent protein (GFP) and immunoglobulin 'Fragment crystallizable' (Fc) chains. Accumulation of the GPs with and without Tm was low, reaching no more than 0.10% of total soluble protein (TSP) in homozygous seed. However, addition of stabilizing domains boosted accumulation up to a maximum of 2.74% of TSP when GFP was used, and albeit less effectively, also the Fc chains of the porcine IgG3 and murine IgG2a increased antigen accumulation, to 0.96% and 1.81% of TSP respectively, while the murine IgG3 Fc chain did not. Antigens with Tm were less susceptible to these manipulations to increase yield. All the antigens were produced in the endoplasmic reticulum and accordingly, they carried high-mannose N-glycans. The immunogenicity of several of those antigens was assessed and we show that vaccination with purified antigen elicited antibodies with virus neutralizing activity in mice but not in pigs.

Introduction

PRRSV is a pig pathogen that emerged in the late eighties in North America and North Europe (Collins et al., 1992; Wensvoort et al., 1991). Since its emergence, PRRSV has spread across the globe. It is considered one of the major pathogens affecting pig industries and causes severe economic losses worldwide (Neumann et al., 2005; Zhou & Yang, 2010). It provokes respiratory distress in pigs of all ages, but is especially problematic when infecting pregnant sows, leading to late abortion, early farrowing and birth of dead or weakened piglets (Terpstra et al., 1991). Lately, more virulent strains with a high incidence of pig mortality have been circulating (Karniychuk et al., 2010; Zhou & Yang, 2010).

PRRSV is an enveloped RNA virus belonging to the family of the *Arteriviridae*. Two genotypes, a North American (NA-type) and a European (EU-type) have been identified, which again can be divided in several subtypes (Meulenberg et al., 1993b; Stadejek et al., 2008). It has multiple open reading frames embedded in its 15 kb long genome (Meulenberg et al., 1997b). Open reading frames

2 to 7 code for the glycoprotein (GP) 2, the E protein, GP3, GP4, GP5, the M protein and the nucleocapsid protein N, respectively (Meulenberg et al., 1995; Meulenberg et al., 1997b; Wu et al., 2001; Wu et al., 2005). Of these structural proteins, the GP5 and the M protein are the major proteins found in the lipid envelope and occur as a disulfide-linked heterodimer (Mardassi et al., 1996). The minor envelope proteins, i.e. GP2, GP3, GP4 and the E protein, are present as non-covalently associated heteromultimers.

PRRSV infection induces a defective immune response with late appearance of neutralizing antibodies and delayed cell mediated immunity (Kimman et al., 2009; Lopez & Osorio, 2004). This allows the infection to be persistent and makes it difficult to eliminate PRRSV from an infected herd. Control of PRRSV is focused on the administration of prophylactic vaccines to minimize the clinical impact of an infection. The vaccines currently used in the field consist of killed and live, attenuated vaccines. Commercial killed vaccines are safe and can provide some protection against homologous viruses, but are totally ineffective against heterologous strains (Geldhof et al., 2012; Zuckermann et al., 2007). The available attenuated vaccines, however, are very effective against homologous challenges and can, to a certain degree, even cross protect against heterologous strains, but they raise safety concerns because of their live viral content (Díaz et al., 2006; Geldhof et al., 2012; Key et al., 2003; Nielsen et al., 2001; Zuckermann et al., 2007). The attenuated vaccine can revert to virulence and cause infection in the vaccinated herd instead of preventing it. This has been documented for PRRSV on several occasions (Key et al., 2003; Nielsen et al., 2001; Opriessnig et al., 2002). There is thus an urgent need for an adaptable, safe and effective PRRSV vaccine, such as a rationally engineered subunit vaccine. The latter is inherently safe because no virus is present and by incorporating elements known to be important in immunity against PRRSV, could be very effective. To this goal, we selected GP3, GP4 and GP5 of the European prototype PRRSV, the 'Lelystad Virus' (LV), to be included in a subunit vaccine. Since viral glycoproteins are complex proteins that cannot be correctly produced in bacteria, we opted to express them in the seed of *Arabidopsis thaliana*, a production platform that has previously been used to generate several heterologous proteins to levels as high as 15% of total soluble protein (TSP) (De Jaeger et al., 2002; Goossens et al., 1999; Loos et al., 2011a; Loos et al., 2011b; Van Droogenbroeck et al., 2007). Production of high-value pharmaceuticals in plants has gained popularity during the last decades, due to intrinsic advantages offered by this production platform, such as flexible scale up, the absence of human or animal pathogens, and the potential for the cheap production and storage of complex proteins (Nagels et al., 2012b; Paul & Ma, 2010; Rybicki, 2010). The technology has reached maturity and several plant made pharmaceuticals are in clinical trials, on the road to the market (Yusibov et al., 2011). Different plant species and tissues have been employed for the production of PRRSV vaccines. Banana leaves

(Chan et al., 2013), potato tubers (Chen & Liu, 2011) and tobacco leaves (Chia et al., 2011; Chia et al., 2010) were used to express the GP5 and recombinant variants thereof. The M protein was produced in corn calli (Hu et al., 2012) and the N protein in soybean seed (Vimolmangkang et al., 2012). Antigen accumulation was relatively low, ranging from 0.011% of TSP for GP5 expressed in tobacco to 0.65% of TSP for N protein expressed in soybean seed and none of the antigens were purified. However, to overcome economic hurdles during vaccine production, it is essential that the antigens accumulate to high levels and that an efficient purification protocol is in place. With this in mind, we manipulated the GP3, GP4 and GP5 coding sequence to augment their production and to facilitate purification by removing transmembrane domains (Tm), adding affinity tags and/or fusing them to stabilizing protein domains, such as the 'Fragment crystallizable' (Fc) of immunoglobulins (IgGs) or the green fluorescent protein (GFP), hence giving rise to an entire set of antigen formats derived from the EU-prototype LV. All these different constructs were cloned in a seed-specific expression cassette (Morandini et al., 2011), transformed into *Arabidopsis thaliana* and the resulting transformants screened for antigen content. The antigens were characterized and evaluated for their ease of production and purification. Further, the immunogenicity of the antigens and their potential to raise neutralizing antibodies were investigated by vaccinating mice, and subsequently piglets, with purified antigens.

This paper details the *in planta* production of a set of antigenic proteins derived from the glycoproteins of the EU-prototype LV. We show that the antigens are correctly produced in the seed, accumulate to levels that are economically feasible (1% of TSP or more; (Rybicki, 2009)) and can be purified by single-step affinity chromatography. Additionally, we assayed the immunogenicity of several antigens in both a murine and porcine model.

Results

Antigen expression in A. thaliana seed

The PRRSV envelope proteins GP3, GP4 and GP5 (Figure 1B) were selected and expressed as different formats in *A. thaliana* seed (Figure 1C). Both full-length GP4 and GP5, as well as their truncated formats without Tm (hereafter referred to as GP4-Tm and GP5-Tm) were cloned into an expression cassette with regulatory sequences of *Phaseolus vulgaris* designed for high seed-specific expression (De Jaeger et al., 2002; Goossens et al., 1999; Morandini et al., 2011) (Figure 1A). All antigens were targeted to the endoplasmatic reticulum (ER) by N-terminal fusion to the signal peptide of a 2S2 seed storage protein, and retained there via a C-terminal KDEL-tag. Little is known about the transmembrane topology of the full-length GP4 and GP5. The C-terminal hydrophobic domain of GP4 is predicted to span the membrane once, whereas the central hydrophobic domain of GP5 is predicted to span the membrane anywhere from one to three times (Meulenberg et al., 1995; Wissink et al., 2005). *In vivo*, transit through the secretory pathway decorates the GPs with complex mammalian N-glycans (Mardassi et al., 1996; Meulenberg et al., 1997a; Meulenberg et al., 1995). However, retaining the GPs in the ER will cause their N-glycans to be of the high-mannose type (Wissink et al., 2005). To evaluate the influence of these ER-associated high-mannose glycans on antigen immunogenicity, a deglycosylated format of the GP4-Tm (named sGP4-Tm) was designed. This was done by mutating the asparagine in the consensus sequence N-x-S/T (with x being any amino acid except proline) into a glutamine for each of the four N-glycan attachment sites that was present. For GP3, only a truncated format without Tm (named GP3-Tm) was cloned. Fusions were also made between these proteins and stable protein domains, such as the GFP and the Fc chains of the porcine IgG3 (pFc), the murine IgG3 (mFc3), and the murine IgG2a (mFc2a), generating a whole set of antigenic PRRSV proteins (Figure 1C). All constructs except the murine Fc fusions, contain the His₆-tag. For every antigen construct, twenty transformants were generated, and the line with the highest antigen content and in which the T-DNA was integrated at a single locus, as determined by segregation analysis (see Material & methods and Supplemental Figures 1 to 3), was retained, selfed and propagated to homozygosity.

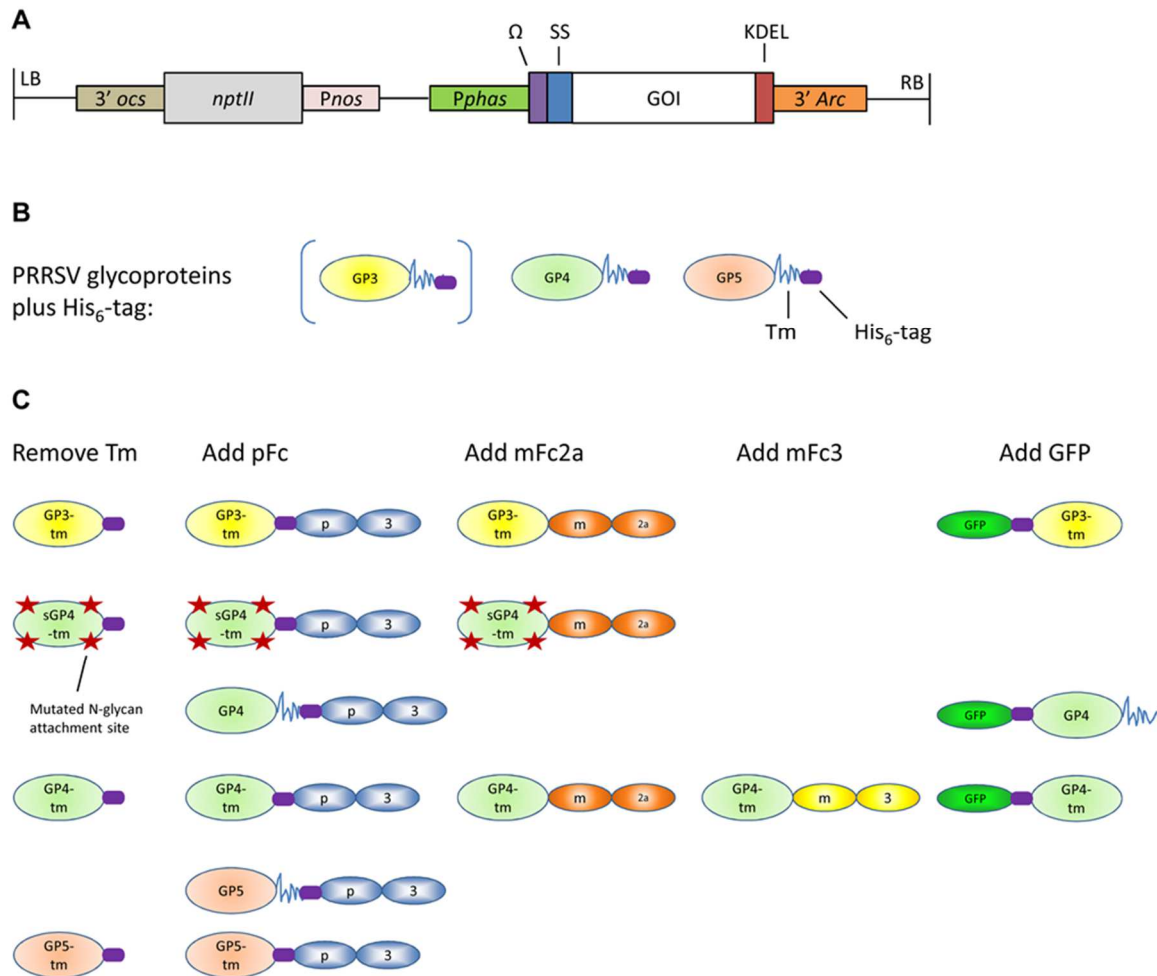


Figure 1. Schematic representation of the T-DNA bearing the seed-specific expression cassette (A), the PRRSV glycoproteins selected to be included in a vaccine (B) and the antigen formats derived thereof (C). (A) The T-DNA carrying the neomycin phosphotransferase II gene which confers resistance to the antibiotic kanamycin and the cassette for high seed-specific expression driven by the β -phaseolin promoter, the leader sequence of the tobacco mosaic virus and the *arcelin5-1* terminator. The cassette is further supplemented with sequences encoding an ER-targeting and retention signal. (B) The full-length PRRSV glycoproteins used in the current study and supplemented with a His₆-tag (purple rod). The jagged line depicts the Tm. The full-length GP3 is placed between brackets as it was never produced as such. Only the GP3 without Tm was expressed, depicted as GP3-Tm in (C). (C) Overview of the manipulations performed on the GPs and the resulting set of antigenic proteins whose coding sequences were cloned into the seed-specific expression cassette. Abbreviations and symbols used: LB and RB, T-DNA left and right border, respectively; 3'ocs, 3' end of the octopine synthase gene; *nptII*, the neomycin phosphotransferase II gene; *Pnos*, promoter of the nopaline synthase gene; *Pphas*, promoter of the β -phaseolin gene; Ω , leader sequence of the 'tobacco mosaic virus'; SS, sequence coding for the signal peptide of the 2S2 *Arabidopsis* seed storage protein; GOI, gene of interest; KDEL, coding sequence of the ER-retention signal; 3'Arc, terminator of the *arceline5-1* gene; Tm, transmembrane domain, jagged line; His₆-tag, immobilized metal affinity chromatography tag, purple rod; pFc, porcine IgG3 Fc, chain of blue ovals; mFc2a, murine IgG2a Fc, chain of orange ovals; mFc3, murine IgG3 Fc, chain of yellow ovals; GFP, green fluorescent protein, bright green oval.

For all the GP3 derived constructs, except the GP3-Tm, transformants could be obtained. Extracts from homozygous GP3-Tm:pFc, GP3-Tm:mFc2a and GFP:GP3-Tm seed stocks were analyzed by western blotting with the anti- (α -) GP3 monoclonal antibody (mAb) VII2D that recognizes the linear epitope MWCKIGHDRCEE (Van Breedam et al., 2011). This revealed several glycoforms migrating at their expected theoretical molecular weight (Mw) ranging from 46.3 kDa (one of the six or seven N-glycan attachment sites occupied) to 60.3 kDa (fully glycosylated) (Figure 2A; Table 1). Mainly for the GFP:GP3-Tm also some non-reduced multimeric structures could be seen between 100 and 150 kDa, together with some degradation products around 37 kDa.

The GP4 seed stocks were analyzed with the α -GP4 mAb XVI11C that recognizes the linear epitope GVSAAQEKISFG (Costers et al., 2010) (Figure 2B). Since GP4 and GP4-Tm carry four N-glycan attachment sites, four different glycoforms could be detected for both, with those of GP4 migrating slower than those of GP4-Tm, due to the presence of the Tm (Figure 2B). However, a differential glycosylation preference was observed, since GP4 showed a higher intensity of its upper band and a very weak signal of the second lowest band, whereas GP4-Tm showed higher intensity for the second band and equal intensities for the other three glycoforms. The signal of sGP4-Tm corresponded with the lower band of GP4-Tm (Figure 2B), confirming that this represents the unglycosylated GP4-Tm. Mutating the N-glycan sites seemed to have a deleterious effect on the accumulation of sGP4-Tm, as shown by the low intensity of its corresponding band. All fusions of GP4, GP4-Tm and sGP4-Tm with pFc, mFc3, mFc2a and GFP migrated at their expected theoretical Mw (Table 1), but with different signal intensities (Figure 2B), pointing toward differences in antigen accumulation in the seed. For GP4-Tm:mFc2a some non-reduced homodimers were visible around 100 kDa together with degraded fragments at 25 kDa (Figure 2B). To a lesser extent, non-reduced multimers and degradation products could also be observed for some of the other antigens (Figure 2B).

Western blot analysis of seed extracts containing the different GP5 antigens and detected with the α -GP5 mAb II5D (which recognizes the linear epitope DSSTYQYIYNLT (Van Breedam et al., 2011)), showed only for GP5-Tm:pFc a weak signal at 50 kDa (Figure 2C, left panel). In all samples however, even in the wild-type extract, a strong background band appeared around 20-25 kDa, possibly covering other signals. A similar pattern was observed when the α -GP5 mAbs II5A and VII2H were used to detect the antigens (data not shown, it should however be noted that the mAbs II5D, II5A and VII2H recognize the same linear epitope (Van Breedam et al., 2011)). Western blot analysis was further performed with an antibody against the His₆-tag, which was also present in the constructs (Figure 1C) and enabled detection of all GP5 antigens (Figure 2C, right panel). As expected, GP5 migrates slower than GP5-Tm. The smaller, faint bands were hypothesized to represent partially or unglycosylated proteins, since both GP5 and GP5-Tm carry two N-glycosylation sites, and this was

confirmed by deglycosylation analyses (see next section). GP5:pFc migrates slower than GP5-Tm:pFc, but is only very faintly detected, indicating low expression. GP5-Tm:pFc on the other hand accumulates notably higher than the other GP5 antigens, but also shows a second band migrating at a size smaller than the theoretical Mw of 49 kDa, which does not correspond with the partially or unglycosylated proteins, but represents a cleavage product (see next section).

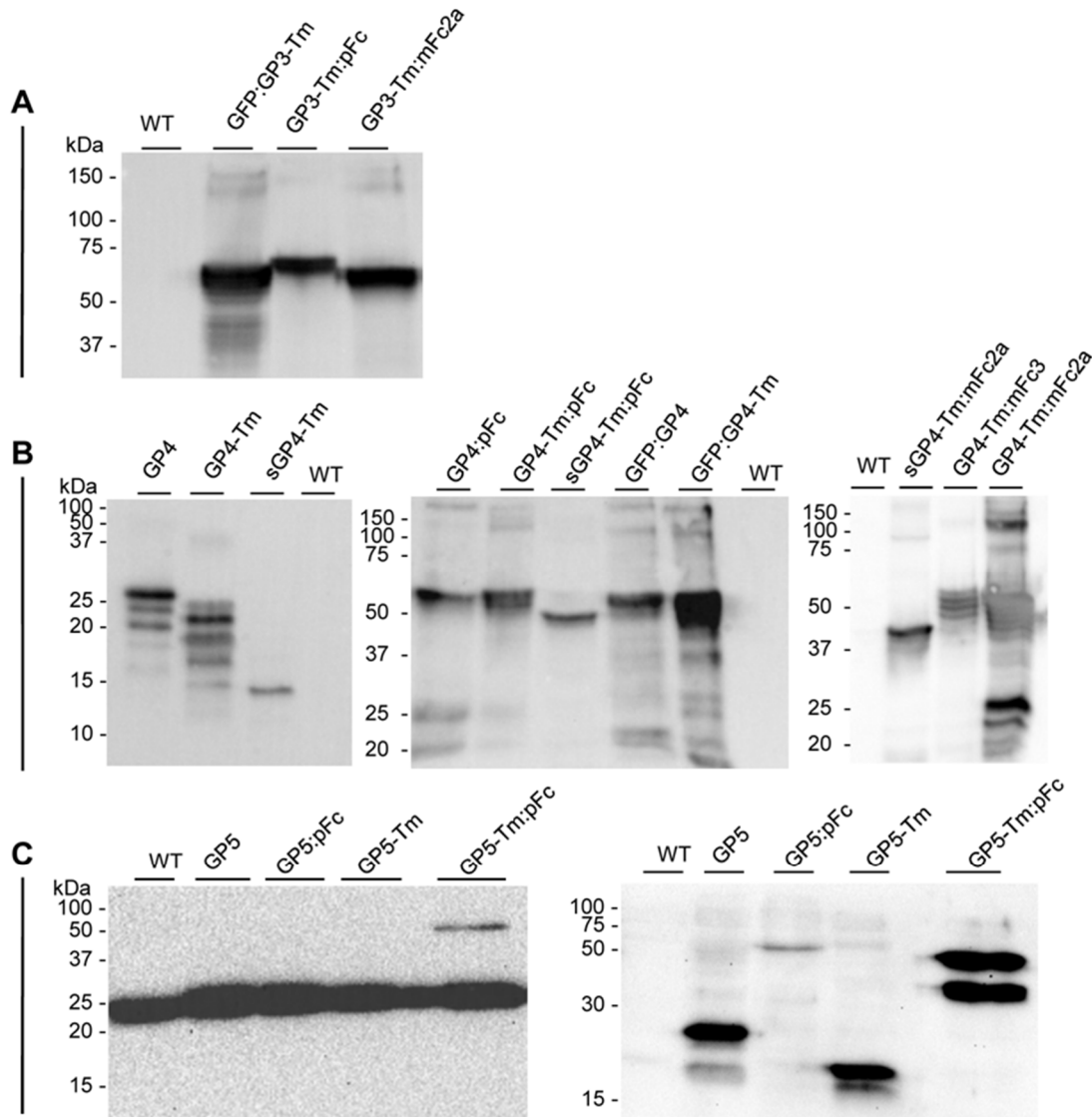


Figure 2. Characterization of the GP3, GP4 and GP5 antigens. Seed extracts of homozygous lines (unless stated differently) producing the different antigen formats were analyzed by western blotting under reducing conditions. Sixty μ g of TSP was loaded. (A) The GP3 antigens detected with the α -GP3 mAb VII2D (Van Breedam et al., 2011). (B) The GP4 antigens detected with the α -GP4 mAb XVI 11C (Costers et al., 2010). For the sGP4-Tm, sGP4-Tm:pFc and the GFP:GP4, segregating T2 stocks were analyzed and for the GFP:GP4-Tm and GP4-Tm:mFc2a, only 15 μ g of TSP was loaded. (C) The GP5 antigens detected with the α -GP5 mAb II5A (Van Breedam et al., 2011) (left panel) or an antibody against the His₆-tag (right panel). For the GP5-Tm:pFc, a T2 segregating stock was analyzed. Abbreviations used: WT, wild-type seed extract; kDa, kilodalton.

In conclusion, all antigens, except the GP3-Tm for which no transformants could be obtained, are present in the seed of the transgenic *A. thaliana* and accumulate to different levels. Also, aside from the GP5-Tm:pFc that is partially cleaved, all antigens appear structurally integer and migrate according to their expected size.

Table 1. Physical parameters of the antigens

Antigen	Amino acids	N-glycosylation sites	Predicted molecular weight (kDa)	
			Unglycosylated	Fully glycosylated
GP3-Tm	166	6	19.2	31.2
GP3-Tm:pFc	407	7	46.3	60.3
GP3-Tm:mFc2a	392	7	44.6	58.6
GFP:GP3-Tm	410	6	46.6	58.6
GP4	176	4	19.5	27.5
GP4:pFc	417	5	46.7	56.7
GFP:GP4	420	4	46.9	54.9
GP4-Tm	158	4	17.6	25.6
GP4-Tm:pFc	399	5	44.7	54.7
GP4-Tm:mFc3	384	6	42.8	54.8
GP4-Tm:mFc2a	384	5	43.0	53.0
GFP:GP4-Tm	402	4	45.0	53.0
sGP4-Tm	158	0	17.6	17.6
sGP4-Tm:pFc	399	1	44.8	46.8
sGP4-Tm:mFc2a	384	1	43.0	45.0
GP5	179	2	19.9	23.9
GP5:pFc	420	3	47.1	53.1
GP5-Tm	119	2	13.6	17.6
GP5-Tm:pFc	381	3	43.1	49.1

Table 1. The molecular weight of the fully glycosylated antigens was obtained by taking the weight of the unglycosylated antigens and adding to that the number of potential N-glycosylation sites multiplied by 2 kDa. Abbreviations used: kDa, kilodalton.

Glycosylation of antigens produced in A. thaliana seed

All antigens were translocated to the ER by an ER-targeting signal derived from the 2S2 albumin seed storage protein of *A. thaliana* (Krebbers et al., 1988) and retained there via a C-terminal KDEL sequence (Gomord et al., 1997a). If the antigens are indeed retained in the ER, they should carry N-glycans of the high-mannose type (Pearse & Hebert, 2010). To verify their glycosylation status, all antigens (18 in total) were deglycosylated by treatment with either endoglycosidase H (endoH) or peptide-N-glycosidase F (PNGaseF), which differ in their substrate specificity. EndoH acts on high-mannose N-glycans only (Maley et al., 1989), whereas PNGaseF processes all N-glycans except core α 1,3-fucosylated structures. In plants, this modification occurs in the late Golgi apparatus and can be found on hybrid and complex N-glycans (Gomord et al., 2010). When the antigen carries

ER-typical high-mannose N-glycans, it will be deglycosylated by both enzymes. However, if the antigen is present in other compartments of the secretory pathway and its glycans undergo maturation into hybrid or complex type glycans, then treatment with endoH and PNGaseF will yield differential deglycosylation, except for α 1,3-fucosylated glycans that will not be processed by either of both glycosidases. However, glycoproteins carrying these structures will be resilient to digestion and hence, can be monitored as well.

All antigens are susceptible to deglycosylation by endoH and PNGaseF and no differences can be discerned between treatment with either glycosidase, indicating that the proteins carry high-mannose N-glycans (Figure 3A, B and C). As hybrid and/or complex glycans have been reported on KDEL-tagged recombinant proteins before (He et al., 2012b; Loos et al., 2011a), it is possible that minute quantities of these are present on the antigens as well and that the mobility shift assay did not reveal these. However, the obtained results clearly identify high-mannose N-glycans as the predominant glycoforms. For GFP:GP3-Tm some degradation was observed (Figure 3A). Figure 3B shows that the banded pattern of both GP4 and GP4-Tm was reduced to a single band corresponding to the fully deglycosylated antigen, confirming that it indeed represents different glycoforms. Figure 3B also shows some non-reduced multimers and degradation for GFP:GP4-Tm and GP4-Tm:mFc2a. Similarly, for the GP5 antigens no differential behavior could be observed when treated with endoH or PNGaseF (Figure 3C). Deglycosylation of GP5 and GP5-Tm showed that the faint bands present before treatment correspond to partially and unglycosylated antigens, since both deglycosylation treatments resulted in the reduction of this pattern to one single band (Figure 3C). Given that the pattern of GP5-Tm:pFc (right panel, Figure 2C and 3C) was not reduced to one single band after endoH or PNGaseF treatment, the lower band does not correspond to the unglycosylated protein, but to a cleavage product (Figure 3C right panel).

To conclude, the antigens were post-translationally modified by N-glycosylation and glycosidase treatment showed that all antigens are decorated with ER-typical high-mannose N-glycans, indicating their correct production in the ER.

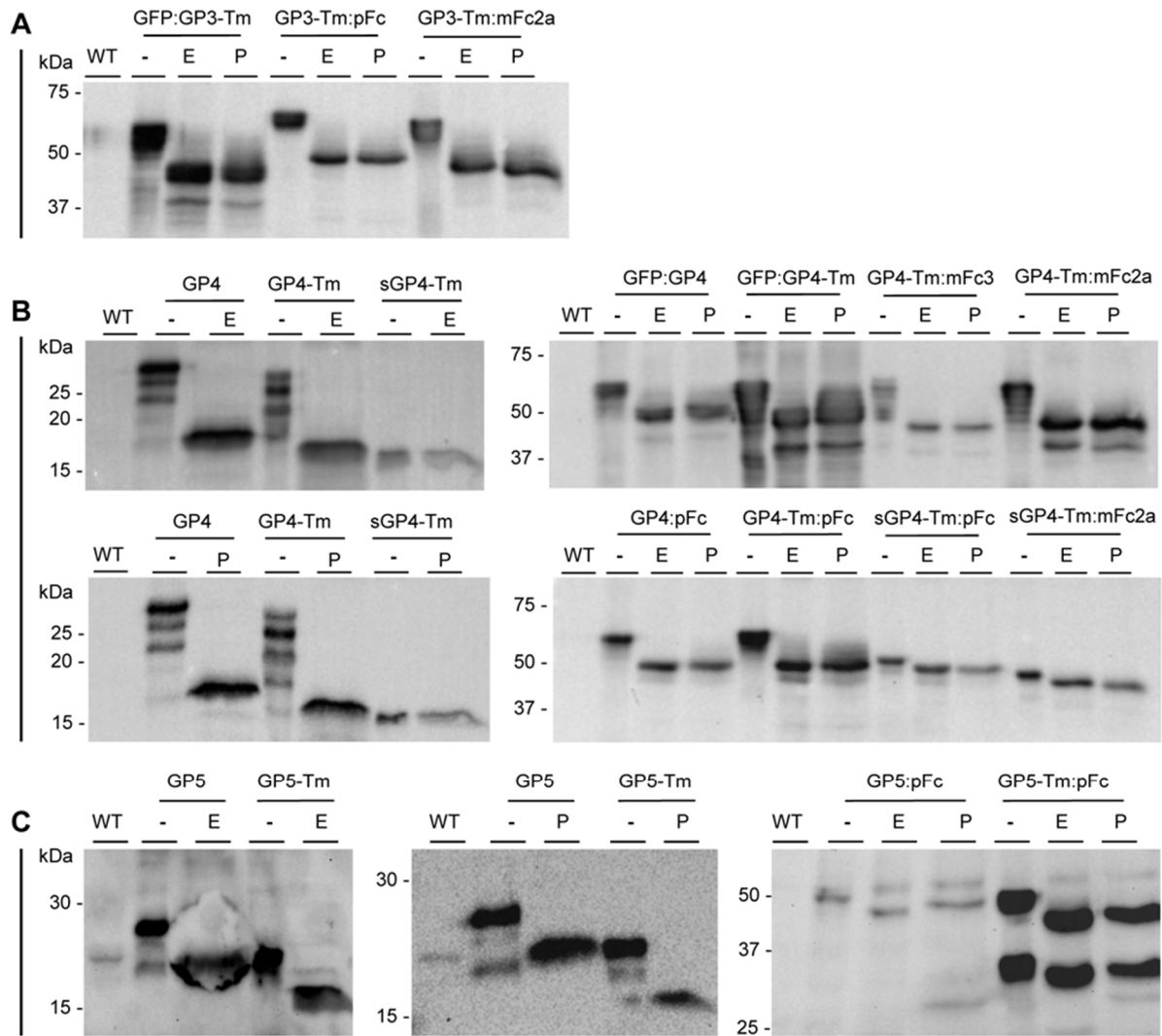


Figure 3. Glycan analysis of the GP3, GP4 and GP5 antigens. Seed extracts from homozygous lines (unless stated differently) producing the different antigen formats were treated with endoH or PNGaseF and analyzed by western blotting under reducing conditions. Thirty μ g of TSP was loaded. (A) Glycan analysis of the GP3 antigens detected with the α -GP3 mAb VII2D (Van Breedam et al., 2011). (B) Glycan analysis of the GP4 antigens detected with the α -GP4 mAb XVI 11c (Costers et al., 2010). For the sGP4-Tm, sGP4-Tm:pFc and GFP:GP4, a segregating T2 stock was used. (C) Glycan analysis of the GP5 antigens detected with an antibody against the His₆-tag. For the GP5-Tm:pFc, a segregating T2 stock was analyzed. Abbreviations used: -, no treatment; E, endoH treatment; P, PNGaseF treatment: WT, non-treated wild-type seed extract; kDa, kilodalton.

Accumulation levels of antigens produced in *A. thaliana* seed

ELISAs were performed to determine the accumulation level of the antigens in the transgenic seed and to assess the effect of fusing them to stabilizing protein domains (Table S1). The antigen content was expressed as antigen mass per gram of dry seed and, after measuring the protein concentration of the seed extracts, as percentage of TSP. For most antigens, T3 homozygous (Ho) lines were analyzed. When not available, the T2 segregating (Se) seed stock was used (Table S1). Therefore, it should be taken into account that, due to a gene dosage effect (De Jaeger et al., 2002), it is possible that the accumulation of the antigens, for which a segregating seed stock was analyzed, could increase when these lines are propagated to homozygosity. During the ELISA experiments, no signal could be obtained for GP5 and GP5:pFc. However, western blotting clearly showed that they were present in seed extracts (Figure 2C, right panel), and thus GP5 and GP5:pFc antigen accumulation was determined by signal quantification after western blot analysis via the His₆-tag (data not shown).

A graphical representation of the accumulation of the antigens (given as percentage of TSP, Figure 4) shows that there was a large variation in accumulation between antigens, ranging from 0.005% of TSP for sGP4-Tm to 2.74% of TSP for GFP:GP3-Tm. Some trends could however be observed. The full-length antigens GP4 and GP5 did not accumulate well (0.08% and 0.10% of TSP, respectively) and removing their Tm to generate GP4-Tm and GP5-Tm did not alter this considerably, but fusion to stabilizing protein domains increased accumulation by several factors, up to 2.36% of TSP for GFP:GP4-Tm. Fusions between GP3-Tm and the same stabilizing domains showed a similar increase, resulting in percentages as those measured for the corresponding GP4-Tm fusions. It is noteworthy that addition of an IgG Fc to the antigens without Tm can have a beneficial effect on their accumulation, as described by De Buck et al. (2013), but that this is highly dependent of the Fc added. Whereas the mFc3 did not enhance GP4-Tm accumulation, the pFc, and the mFc2a even more so, notably did so, boosting accumulation from 0.05% to 0.67% and 1.59% of TSP, respectively. The full-length antigens appeared to be less prone to manipulation by addition of stabilizing domains. In contrast to GP4, where addition of the pFc and GFP had little impact on accumulation (from 0.08% to 0.14% and 0.15% of TSP, respectively), coupling of the pFc to GP5 actually reduced accumulation of the resulting GP5:pFc (from 0.10% to 0.01% of TSP). Removing the four N-glycan attachment sites of GP4-Tm in sGP4-Tm was also deleterious, since accumulation was diminished from 0.05% to 0.005% of TSP. Again, addition of the pFc and mFc2a to sGP4-Tm slightly increased accumulation. Far outperforming any other format, the N-terminal GFP fusions, i.e. GFP:GP3-Tm and GFP:GP4-Tm, accumulated to 2.74% and 2.36% of TSP, respectively.

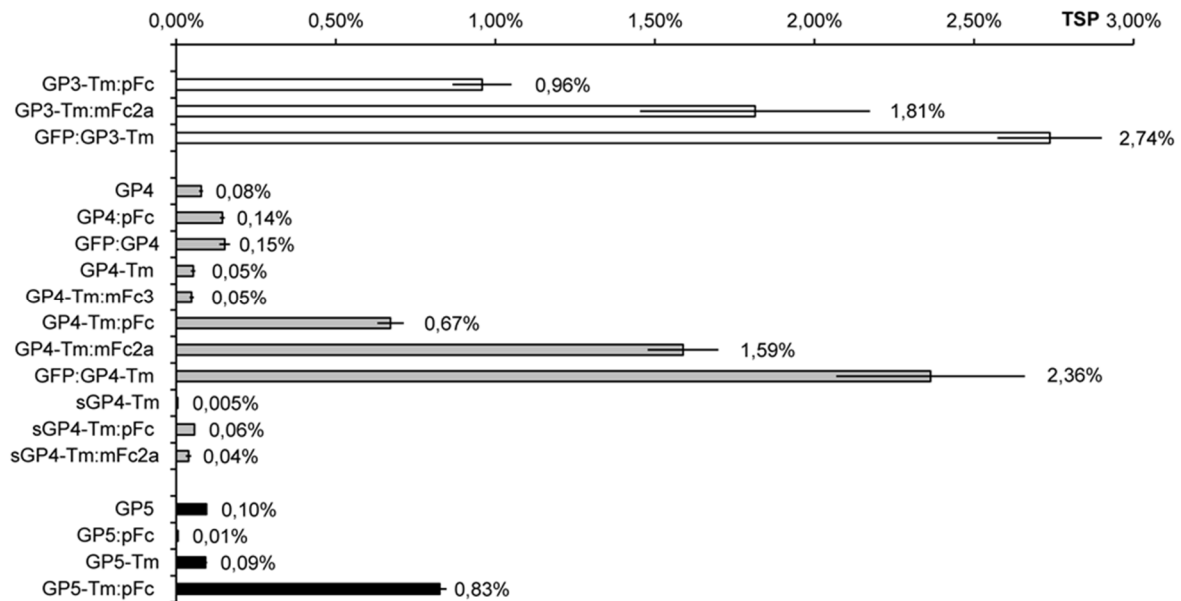


Figure 4. Antigen accumulation in transgenic seeds as determined by ELISA and given as percentage of TSP. Values are the mean of three biochemical repeats, error bars represent the standard deviation. Accumulation of the GP3, GP4 and GP5 antigens is depicted as white, grey and black bars, respectively. Abbreviation used: TSP, total soluble protein.

Purification of antigens produced in *A. thaliana* seed

Several techniques can be used for purification. Antigens fused to an IgG Fc are suitable for protein-A affinity chromatography, whereas the formats carrying a His₆-tag can be purified by immobilized metal affinity chromatography (IMAC).

A high and low accumulating antigen, respectively GFP:GP4-Tm and GP4-Tm, were selected to be purified by IMAC. The IMAC purification of GP4-Tm resulted in fractions enriched with GP4-Tm as observed by western blot analysis (the banded pattern corresponds to the different glycoforms of GP4-Tm, Figure 5B), but GP4-Tm co-eluted with contaminating proteins as shown on the Coomassie-stained gel (Figure 5A). IMAC was thus efficient in capturing GP4-Tm, however, since the eluted fractions were only semi-pure, additional steps are necessary to obtain pure GP4-Tm. No degradation of the GP4-Tm was visible (data not shown). In contrast to GP4-Tm, the eluted GFP:GP4-Tm was of high purity (80 – 82%, Figure 5C). Western blot analysis (Figure 5D) showed that some GFP:GP4-Tm was still present in the flow through and after washing to remove imidazole, the final yield was only 40%. Nonetheless, successive rounds of loading the flow through onto the column and subsequent elution could increase the recovery up to 75% (data not shown). Also, small quantities of degraded GFP:GP4-Tm were visible both in the crude extract as in the purified samples

(Figure 5D). These experiments suggest that when IMAC is used to purify proteins from crude *A. thaliana* seed extracts, some endogenous proteins co-purify and that the level of contaminating protein depends on the amount of target protein present in the extract. When the target protein is abundant, then it is preferentially retained by the column. Low abundant proteins on the other hand are co-purified with higher levels of endogenous proteins. In summary, IMAC is suitable to capture GP4-Tm, but additional steps are necessary to obtain pure protein. Considering GFP:GP4-Tm and taking a yield penalty into account, IMAC can be used to directly purify the antigen.

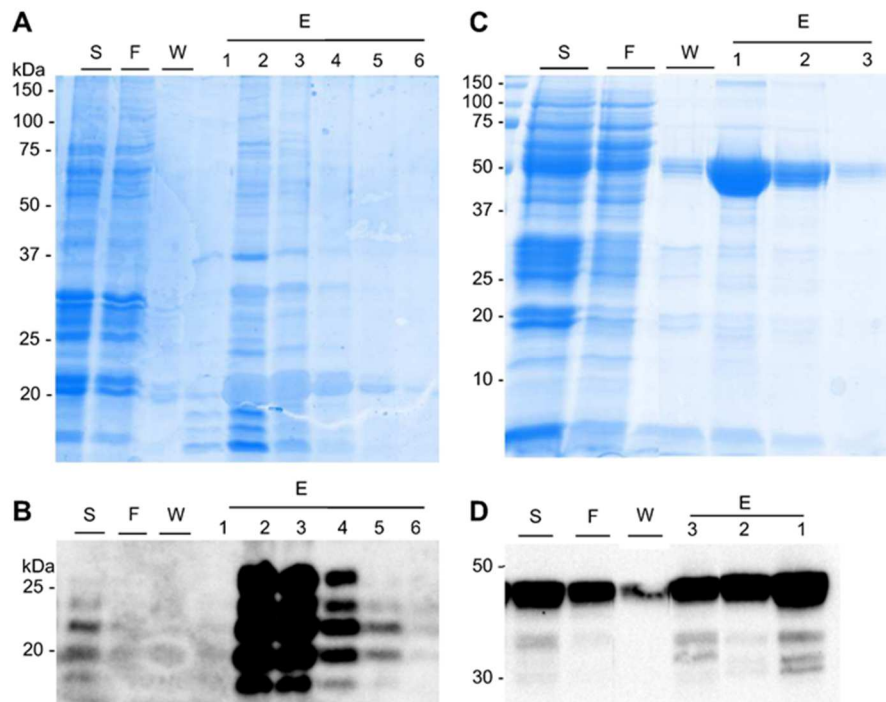


Figure 5. IMAC purification of selected antigens. One gram of homozygous seed stocks containing 0.14 mg GP4-Tm or 7.37 mg of GFP:GP4-Tm was extracted and subjected to IMAC. Purification of the GP4-Tm was analyzed by 12% SDS-PAGE and Coomassie Blue staining (A) or western blotting (B). Purification of the GFP:GP4-Tm was analyzed by SDS-PAGE (15%) and Coomassie Blue staining (C) or western blotting (D). All gels were run under reducing conditions and an antibody against the His₆-tag was used to detect the antigens during western blot analysis. Abbreviations used: S, sample; F, flow through; W, wash; E, elution; kDa, kilodalton.

The GP4-Tm antigens fused to pFc, mFc3 and mFc2a were selected to be purified by protein-A affinity chromatography (Figure 6). The strong and specific interaction between protein-A derived from *Staphylococcus aureus* and the Fc part of the antigens should allow for efficient purification (Goding, 1978). Coomassie and western blot analysis showed that all three Fcs used had affinity for the

protein-A resin and that both low (GP4-Tm:mFc3 in Figure 6C) and high accumulating antigens (GP4-Tm:pFc and GP4-Tm:mFc2a, Figure 6A and B, respectively) could be captured and purified by single-step protein-A chromatography. The recovery and purity for GP4-Tm:pFc were 55 and 85%, respectively, and for GP4-Tm:mFc2a 60 and 88%, respectively (they were not determined for GP4-Tm:mFc3).

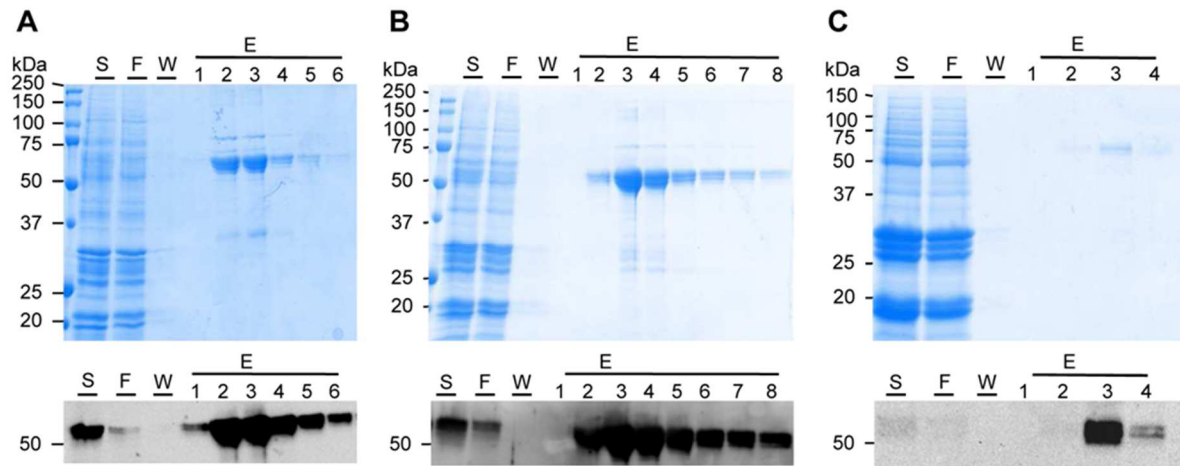


Figure 6. Protein-A affinity chromatography of selected antigens. For each antigen, one gram of homozygous seeds (equivalent to 0.15 mg GP4-Tm:mFc3, 2.06 mg GP4-Tm:pFc and 5.15 mg GP4-Tm:mFc2a) was extracted and subjected to protein-A chromatography. Purification of GP4-Tm:pFc (A), GP4-Tm:mFc2a (B) and GP4-Tm:mFc3 (C) was analyzed by 12% SDS-PAGE and Coomassie Blue staining (upper panels) or western blotting (lower panels). All gels were run under reducing conditions and the α -GP4 mAb XVI 11C was used to detect the antigens during western blotting. Abbreviations used: S, sample; F, flow through; W, wash; E, elution; kDa, kilodalton.

Immunogenicity of antigens produced in *A. thaliana* seeds

To study the immunogenicity of the plant-produced antigens, the GFP:GP4-Tm, GP4-Tm:mFc2a and sGP4-Tm:mFc2a, all carrying the GP4-Tm, were selected to vaccinate mice. The GP4-Tm was chosen as a model antigen as it contains a B-cell epitope that consistently induces neutralizing antibodies upon infection and neutralizing antibodies are considered a good correlate for protection against PRRSV (Lopez & Osorio, 2004; Meulenberg et al., 1997a). Mice were injected with 5 μ g of purified antigen mixed with an oil-in-water adjuvant following a prime-boost protocol with three weeks between both vaccinations. Figure 7A shows that after six weeks high titers of GP4-Tm-specific antibodies were induced by all three antigens as measured by ELISA. The sera collected at week six were then subjected to an immuno peroxidase cell monolayer assay (IPMA) to check for viral antigen

recognition (Figure 7B). Despite the high antibody titers induced by all antigens, only GP4-Tm:mFc2a generated antibodies capable of PRRSV recognition in IPMA. Subsequently these sera were shown to have virus neutralizing capacity in a seroneutralization test (Figure 7C). All three antigens thus raised high levels of antigen-specific antibodies in a mouse model, and those of the GP4-Tm:mFc2a vaccinated group were also able to neutralize the PRRSV.

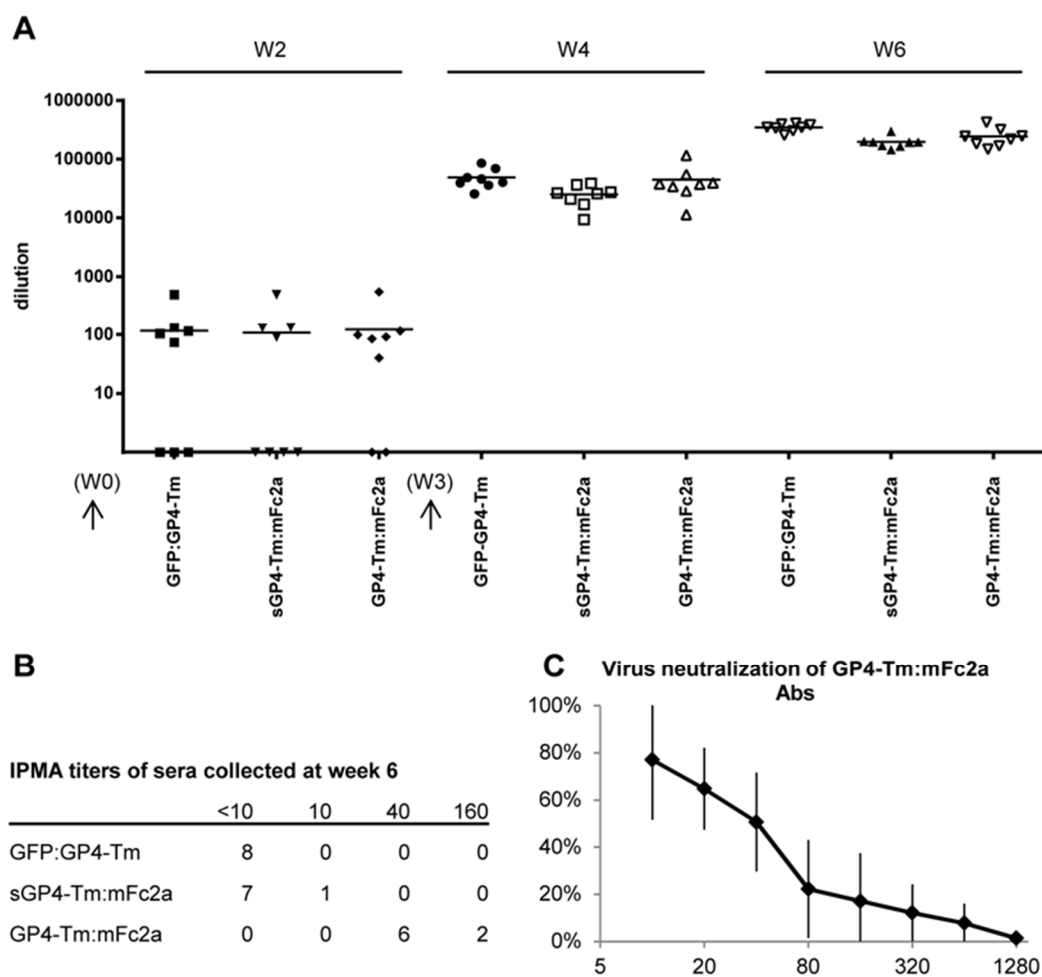


Figure 7. Characterization of the immunogenicity of selected GP4 antigens in a mouse model. Per antigen, eight mice were vaccinated twice with 5 μ g antigen mixed with an oil-in-water adjuvant. The booster vaccination was administered at week three (both vaccinations are depicted as black arrows). (A) Antigen-specific IgG titers induced by the GFP:GP4-Tm, GP4-Tm:mFc2a and sGP4-Tm:mFc2a at weeks two, four, and six after vaccination as measured by ELISA. Per vaccinated group, the individual values for each mouse and the mean (horizontal line) are given. (B) The titers of the sera collected at week six when tested via IPMA. Per vaccinated group the number of mice with a given titer is indicated. (C) The sera at week six of the GP4-Tm:mFc2a vaccinated mice were subjected to a seroneutralization assay to determine their virus neutralizing activity. The mean percentage of virus reduction (n = eight mice) in function of the dilution is given, error bars represent the standard deviation.

Upon completion of the murine vaccination trial, a second study was initiated to assess the immunogenicity of antigens carrying the pFc in pigs, the *in vivo* host of the PRRSV. During this preliminary trial three piglets were vaccinated with an antigen cocktail consisting of 100 µg purified GP3-Tm:pFc, 100 µg purified GP4-Tm:pFc and 100 µg purified GP5-Tm:pFc mixed with an oil-in-water adjuvant following a prime-boost protocol with a four-week interval between both vaccinations. High GP3-Tm, GP4-Tm and GP5-Tm-specific antibody titers were obtained after the booster vaccination from week four onward (measured by ELISA; Figure 8A, B and C). When sera were tested via IPMA, high titers were found for one piglet from week five onward, however, no response could be detected for the other two piglets (Figure 8D). Sera from all three piglets collected at week five and six were subjected to a seroneutralization test, but no response above the threshold could be detected (data not shown).

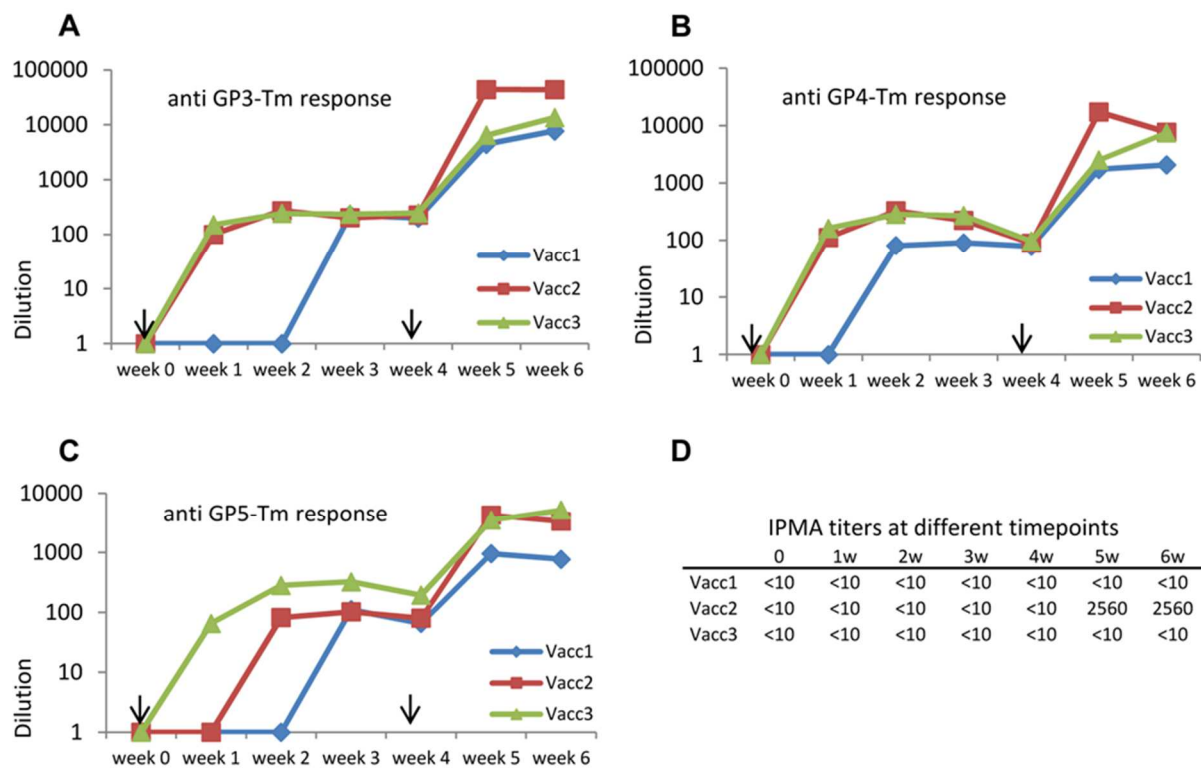


Figure 8. Characterization of the immunogenicity of selected antigens carrying the pFc in a porcine model. Three piglets were vaccinated twice with an antigen cocktail consisting of purified GP3-Tm:pFc, GP4-Tm:pFc and GP5-Tm:pFc mixed with an oil-in-water adjuvant. The booster vaccination was administered at week four (both vaccinations are depicted as black arrows). At different time points, the titers of IgG antibodies specific for GP3-Tm (A), GP4-Tm (B) and GP5-Tm (C) was determined. The collected sera were also tested via IPMA and titers of sera collected from week zero to six given (D).

Chapter II

In summary, all antigens vaccinated in mice and piglets generated high titers of antigen-specific antibodies after the booster vaccination and the sera of mice vaccinated with GP4-Tm:mFc2a also had virus neutralizing capacity. In contrast to this, no such activity could be detected in any of the sera of the vaccinated piglets.

Discussion

We designed, cloned and successfully expressed a set of antigenic proteins derived from the glycoproteins of the PRRSV EU-prototype LV in the seed of *A. thaliana*. Analysis of antigen glycosylation status showed that the antigens were post-translationally modified by N-glycosylation, and that they carried high-mannose N-glycans, corresponding to their production in the ER and possibly, also indicating their retention there. For GP4 this scenario would mean that its single-span transmembrane domain (Meulenberg et al., 1995), predicted to leave the KDEL-tag in the cytosol out of reach of the KDEL receptor, is not correctly inserted into the ER membrane or, that GP4 is retained in the ER in a KDEL-independent manner. In animal cells, GP4 is withheld in the ER until it associates with GP2 and GP3 to form a heterotrimer after which the complex is released (Wissink et al., 2005) and a similar mechanism might function in the seed. For GP5, the exact topology of the Tm, which is predicted to span the ER membrane anywhere from one to three times (Meulenberg et al., 1995), is not known, but the presence of high-mannose glycans would strengthen the hypothesis that the KDEL-tag is functional.

On the other hand, we cannot exclude that the antigens have migrated to other subcellular compartments, since recombinant proteins with ER-typical glycosylation have been found in non-ER compartments before in the seed of Arabidopsis. Using the same expression vector, the anti-HIV monoclonal antibody 2G12 was found almost exclusively in protein storage vacuoles (Loos et al., 2011a), whereas a synthetic antibody based on the 2G12 was deposited in newly formed ER-derived vesicles (Loos et al., 2011b) and another synthetic antibody even in the periplasmic space (Van Droogenbroeck et al., 2007). All these recombinant proteins carried the KDEL sequence and were decorated to some extent with high-mannose glycans, showing that the mere presence of these glycans is not sufficient to allocate a protein to the ER. Further localization studies are thus needed to pinpoint the exact subcellular compartment in which the antigens reside.

Despite the advantages offered by the production of recombinant proteins in plants, such as safety, flexible scalability and complex protein synthesis, low protein yields have been a major problem during previous studies concerning PRRSV antigens (Chan et al., 2013; Chen & Liu, 2011; Chia et al., 2011; Chia et al., 2010; Hu et al., 2012; Vimolmangkang et al., 2012). Antigen accumulation ranged from 0.011% of TSP for recombinant GP5 expressed transiently in tobacco leaves (Chia et al., 2010) to 0.65% of TSP for the N protein produced in soybean seed (Vimolmangkang et al., 2012). To overcome this hurdle, we included in our set of antigenic proteins a number of antigen formats carrying stable protein domains, such as GFP or an immunoglobulin Fc fragment, which are expected to increase antigen accumulation. Indeed, a 6-fold increase (taking into account the similar molecular

weight of GP4-Tm and the pFc domain) in accumulation of the GP4-Tm component was observed when comparing the GP4-Tm antigen (0.05% of TSP) to the GP4-Tm:pFc (0.67% of TSP), rising to over 15-fold for the GP4-Tm:mFc2a (1.59% of TSP) and even more than 20-fold when comparing to the GFP:GP4-Tm (2.36% of TSP). Considering the GP5 antigens, both GP5 and GP5-Tm accumulated to about 0.09% of TSP and this rose to 0.83% of TSP for GP5-Tm:pFc, corresponding to 2.3 mg/g of seed weight. These numbers are considerably higher than those obtained for recombinant GP5 produced in banana (0.037% of leaf TSP (Chan et al., 2013)), potato (4.7µg/g of leaf fresh weight (Chen & Liu, 2011)), tobacco (0.011% of TSP (Chia et al., 2010)) or for a fusion between GP5 and the B subunit of *Escherichia coli* heat-labile enterotoxin produced in tobacco (0.015% of TSP (Chia et al., 2011)).

Further, our results are in line with those obtained by Obregon et al. (2006) who fused the HIV-1 p24 antigenic protein to the Fc of human IgA, resulting in a 13-fold increase in antigen accumulation and those of De Buck et al. (2013) who could boost the accumulation of Nanobodies® (antigen-binding domains of camelid heavy-chain only antibodies) in the seed of *Arabidopsis* by 10- to 100-fold up to 16% of TSP after fusion with Fcs of different origin. As a general tendency we observed that adding stabilizing domains to the truncated antigens GP3-Tm, GP4-Tm and GP5-Tm had a positive influence on antigen accumulation and that the extent of this influence depended on the domain added, with GFP giving the highest accumulation followed by, in decreasing order, the mFc2a, pFc and mFc3.

GFP has a long history in the field of cell biology and GFP-fusion proteins have been made for decades (Cubitt et al., 1995). The main goal of these fusions has always been to create reporter proteins that could be followed at the cellular level. However, our data indicate that GFP can also effectively be used to augment target protein accumulation and, taken into account its stability and the fact that several GFP-based purification methods have been published (Peckham et al., 2006; Rothbauer et al., 2008), fusion to GFP could represent an alternative strategy to boost target protein production. However, caution should be taken in the context of vaccination where GFP might interfere with the immune response raised against the GFP-tagged antigen.

Manipulation of the full-length antigens GP4 and GP5 by addition of stable protein domains was much less effective, indicating that the Tm somehow compromises antigen accumulation. Mutating the N-glycan attachment sites of the GP4-Tm profoundly disturbed accumulation of the resulting deglycosylated sGP4-Tm. Both sGP4-Tm, as well all fusions herewith, underperformed when compared to the corresponding glycosylated formats. It will be interesting to determine whether this is due to the change in primary structure after amino acid substitution or because of the absence of N-glycans, possibly resulting in a perturbation of interaction with chaperones and the quality control mechanism in the ER. In summary, we observed that stable protein domains can be added to the

antigens as a strategy to increase antigen accumulation with GFP being most beneficial, that the effect of such fusions depends on both the antigen and the stable domain, and that for each antigen the optimal manipulation should therefore be identified on a case by case basis.

In addition to boosting accumulation, the Fc domains also allow for protein-A affinity chromatography. We showed that antigens carrying the pFc, mFc3 and mFc2a and varying in accumulation between 0.15 and 5.15 mg per gram dry seed could all be purified by single-step protein-A affinity chromatography with a recovery of about 60%. On the other hand, single-step IMAC purification of antigen formats with a His₆-tag was restricted to high abundant antigens such as GFP:GP4-Tm. It seems that some seed proteins also show affinity for the resin and compete for binding sites on the IMAC column. If the antigen is low abundant, then more endogenous proteins will bind and co-elute, lowering the final antigen purity and increasing the need for an additional purification step.

The platform to be chosen for subunit vaccine production should allow for high production and correct folding and assembly of the antigens. Regarding the latter, it is doubtful that complex PRRSV antigens can be produced correctly in bacteria. Indeed, disease enhancement was observed when recombinant GP5 from *E. coli* was vaccinated (Prieto et al., 2011). PRRSV antigens were also produced in baculovirus infected insect cells, but only low yields were achieved (Duran et al., 1997). Virus-like particles containing GP5 or GP5 and M could be isolated from the insect cell culture medium, but no yield parameter was reported (Nam et al., 2013; Wang et al., 2012). Also when PRRSV antigens were produced *in planta*, accumulation was rather low, with a maximum of 0.65% of TSP when the N protein was produced in soybean seeds (Vimolmangkang et al., 2012). In contrast to previous studies, we managed to produce several PRRSV-derived antigens to levels of 1% of TSP and higher, a level considered to be economically relevant (Rybacki, 2009), and extracted these to high purity. Subsequently, the immunogenicity of the plant-produced antigens was assessed in a murine model. The GFP:GP4-Tm, GP4-Tm:mFc2a and sGP4-Tm:mFc2a were vaccinated in mice and although all antigens generated high titers of antigen-specific antibodies, only those induced by GP4-Tm:mFc2a were virus neutralizing. Whereas both GFP and mFc2a served to increase antigen accumulation in the seed, the latter also seemed to function as an immune enhancing tag, possibly by targeting the antigen to antigen-presenting cells via interaction with Fc receptors (Czajkowsky et al., 2012; Ye et al., 2011). Deletion of the N-glycans by amino acid substitution was detrimental for antigen immunogenicity, as the sGP4-Tm:mFc2a did not induce virus neutralizing antibodies. The ER-associated high-mannose glycans thus seem beneficial for antigen immunogenicity. Indeed, mannosylation of antigens has been shown to selectively target antigen presenting cells thereby

enhancing antigen immunogenicity (Engering et al., 1997; Tan et al., 1997). In short, GP4-derived antigen is capable of eliciting virus neutralizing antibodies in a murine model, confirming the immunogenicity of the plant-produced antigens. Upon completion of the murine trial, antigen immunogenicity was also preliminary assessed in a porcine model. Three piglets receiving an antigen cocktail consisting of purified GP3-Tm:pFc, GP4-Tm:pFc and GP5-Tm:pFc developed high titers of antigen-specific antibodies, but only the sera of one piglet recognized the PRRSV in the IPMA and no virus neutralizing activity could be detected. In this respect it is interesting to mention the findings of Plana-Durán et al. (1997) who observed that pigs having received an inactivated PRRSV vaccine were partially protected against reproductive failure even though their sera were unresponsive in IPMA before challenge. It has also been shown that some experimental, inactivated PRRSV vaccines confer partial protection even without the presence of virus neutralizing antibodies prior to challenge (Vanhee et al., 2009). Due to a priming effect these vaccinated piglets generate a faster and stronger neutralizing antibody response after challenge. Taken these considerations into account, it becomes clear that the antigens need to be further examined via a challenge experiment to study their capacity to prime the antibody response, stimulate a cellular response and reduce viremia. This will allow a final assessment of the protective capacity of a plant-produced PRRSV vaccine.

In conclusion, we generated a set of antigenic proteins derived from the PRRSV, incorporating elements known to be important in immunity against PRRSV, in *A. thaliana* seed. The antigens were correctly produced in the ER, accumulated to levels that are economically relevant and could be purified by single-step affinity chromatography. The immunogenicity of the plant-produced antigens was confirmed in a murine model and preliminary tested in a porcine model. It is now time to perform a challenge experiment to assess the protective capacity of the plant-produced antigens and continue on the road towards a safe and effective PRRSV vaccine.

Material and methods

Ethics Statement

All animal experiments were approved by the Local Ethical Committee of the Faculty of Veterinary Medicine, Ghent University.

Cloning and generation of transgenic lines

The original sequences of the GP3, GP4 and GP5 genes were derived from the complete LV-genome (GenBank accession number M96262.2) and all the cloned genes described in the following paragraphs were deposited at GenBank (accession numbers KF983786 to KF983804).

All the sequences coding for the antigens containing GP4, GP5 and their truncated counterparts GP4-Tm, sGP4-Tm, and GP5-Tm were cloned in the seed-specific expression cassette of the gateway compatible pPphasGW destination vector (Morandini et al., 2011) containing the β -*phaesolin* promoter, the *arcelin-5I* terminator and the *nptII* gene for kanamycin resistance (De Jaeger et al., 2002). The sequences of the antigens were further complemented at their 5' end with the Kozak motif (CCACC) plus the coding sequence of the ER-targeting signal peptide from the *Arabidopsis* 2S2 seed storage protein and at their 3' end with the coding sequence of the KDEL ER-retention signal. The genes of the GP4, GP5 and truncated formats GP3-Tm, GP4-Tm, sGP4-Tm and GP5-Tm were chemically synthesized flanked by the gateway recombination sites attB1-B2. Additionally, a sequence coding for six histidines (His₆-tag) was inserted upstream of the KDEL coding sequence. A BP recombination reaction with the pDON221 and LR reaction with the pPhasGW allowed transfer of all genes to the destination vector, except those coding for the GP3-Tm and fusions thereof. The latter were cloned into a multisite gateway destination vector using the aforementioned regulatory elements and complemented with the same sequences encoding the ER-targeting and retention signal, resulting in an expression vector bearing the *bar* gene as resistance marker (all Gateway reactions were performed according to the Gateway manual protocol). The codon use of all cloned sequences was optimized to resemble that displayed by *Arabidopsis* seed storage proteins and all genes were sequence verified before transfer to the destination vectors.

Coupling of stabilizing domains to the GP4, GP5 and truncated formats GP3-Tm, GP4-Tm, sGP4-Tm and GP5-Tm was done via classical cloning of the corresponding coding sequences. The sequence of the porcine IgG3 Fc (accession no: EU372658; pFc) was added to that of the antigens using a Gateway compatible pDON221 vector that contained the gene of a camel-like antibody (a fusion between the antigen-binding domain of a heavy-chain camel antibody with an Fc fragment of choice) carrying the porcine IgG3 Fc (Virdi et al., 2013). The sequence of the antigen binding domain was swapped by

restriction digest with the enzymes *EcoRI* and *AvrII* and replaced by that of the desired antigen. To this goal the sequences of the GP4, GP5, and truncated counterparts, were amplified by polymerase chain reaction (PCR) to include the *EcoRI* and *AvrII* restriction sites (Table S2 and S3). An LR reaction transferred the resulting genes to the pPhasGW.

A similar strategy was followed to incorporate the sequences of the murine IgG3 Fc (NCBI accession no X00915; UniProt accession no P03987; mFc3) and IgG2a Fc (derived from the full-length murine IgG2a heavy chain; NCBI accession no AAB59660.1). The sequences of the antigen binding domains of camel-like antibodies with murine IgG3 Fc and IgG2a Fc were swapped by restriction digest with *EcoRI* and *SacII* and substituted by those of selected antigens that had been amplified by PCR (Table S2 and S3). In contrast however to the cloning of the porcine IgG3 Fc, the sequence coding for the His₆-tag was no longer included in the final product. Subsequently, genes were transferred to the destination vectors by LR reaction.

Concerning the GFP fusions, the sequence of the enhanced green fluorescent protein (derived from the pFS7(GW)(Karimi et al., 2002)) was extended *in silico* at its 5' end with the Kozak and 2S2 sequence. At its 3' end the coding sequences of the his₆-tag, the GP4-Tm plus KDEL and a *BamHI* restriction site were added. The resulting GFP:GP4-Tm gene was embedded between the attL1-2 sites and chemically synthesized. By PCR it was extended to form the GFP:GP4 gene (Table S2 and S3). The GP4-Tm sequence in the GFP:GP4-Tm gene was swapped by restriction digest with *SacII* and *BamHI* for an amplicon of the GP3-Tm generating the GFP:GP3-Tm gene (Table S1 and S2). LR reaction transferred all genes to the destination vectors.

Heat shock delivered all expression vectors to *Agrobacterium* strain C58C1Rif^R containing the pMP90 virulence plasmid (Koncz & Schell, 1986) and the resulting strains were used to transform *Arabidopsis thaliana* (L.) Heyhn, ecotype Columbia 0, by floral dip (De Buck et al., 2012). Transformed seeds containing the T-DNA were identified and lines propagated to homozygosity as detailed in (De Buck et al., 2012). In short, T1 seeds obtained after floral dip were grown on selective Murashige and Skoog medium for three to four weeks at 21 °C on a 16-h light / 8-h dark cycle. Twenty transformants were then transferred to soil and grown under the same conditions. These were selfed and when seed setting was complete, transferred to a room at 25 °C and low humidity until completely dry. The dry seeds of each transformant were harvested to obtain twenty T2 seed stocks. These stocks were screened for the presence of antigen by western blot analysis as detailed below and seven stocks with high antigen accumulation were analyzed for the number of integrated T-DNA loci. As the T1 transformants are per definition hemizygous, the T-DNA locus number can be determined by the segregation ratio in the T2 generation. To identify the single locus T-DNA integration lines, seeds of

segregating T2 stocks were sown on selective medium and the observed ratio of germinated over sensitive seeds was compared to the expected 3:1 segregation ratio for single locus lines using a χ^2 -statistical test (De Buck et al., 2012). For each antigen the T2 stock with the highest antigen content and the T-DNA integrated at a single locus in the genome was retained and propagated to homozygosity.

Protein extraction, protein concentration determination, SDS-PAGE, Coomassie staining and western blot analysis

To characterize the antigen production in transgenic seed, TSP was extracted as described elsewhere (De Buck et al., 2012) with minor modifications: 1 ml of cold extraction buffer (50 mM Phosphate buffer, pH 7.8, 0.3 M NaCl, 0.1% (w/v) 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate (CHAPS) supplemented with an ethylenediaminetetraacetic acid (EDTA) free, proteinase inhibitor cocktail (1 tablet per 50 ml; Roche)) was added to 10 mg of liquid nitrogen (N₂) frozen, crushed seeds. Samples were vortexed for 1 minute and centrifuged at approximately 20,000 g and 4°C for 30 minutes. 700 µl supernatant was removed, snap-frozen in liquid N₂ and stored at -20°C. Total protein concentration was determined by the Lowry method as described in (De Buck et al., 2012).

Seed extracts containing TSP were separated by reducing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). To visualize proteins by Coomassie staining, the gels were submerged in InstantBlue™ (Expedeon) solution for an hour and then washed three times for 20 minutes with double distilled water (ddH₂O). Stained gels were photographed with the Molecular Imager® ChemiDoc™ XRS+ Imaging System (Biorad). For western blot analysis the proteins were transferred electrophoretically from a polyacrylamide gel onto Immobilon-P polyvinylidene fluoride membrane (Millipore, Billerica, MA). Blots were blocked in 4% (w/v) skimmed milk and 0.1% (v/v) Tween 20 in phosphate-buffered saline (PBS). Several commercial secondary antibodies coupled to horse radish peroxidase (HRP) were used to develop the western blots. Anti-mouse IgG HRP-linked whole antibody (from sheep, GE healthcare) was used in combination with the α -GP4 mAb XVI11C and α -GP5 mAbs II5D, II5A and VII2H at a dilution of 1:5000 in blocking buffer. The α -GP4 mAb XVIII5G and α -GP3 mAb VII2D were detected with α -mouse IgG1 (gamma-chain)-HRP (from rabbit, Sigma Aldrich) diluted 1:5000 in blocking buffer. As an antibody against the His₆-tag, a penta-His HRP conjugate (Qiagen) was used diluted 1:2000 in blocking buffer. Western blotting was further performed as described in (De Buck et al., 2012) and the released chemiluminescent signal measured with the

abovementioned imaging system. Quantification of signal intensities was performed with the Image Lab 3.0 software (Biorad).

Glycan analysis

Seed extracts containing TSP were treated with peptide N-glycosidase F (PNGaseF, New England Biolabs) and recombinant endoglycosidase H (endoH, New England Biolabs) to analyze antigen glycosylation status. Glycosidase reactions were performed according to the manufacturer's protocol. In short, to 9 μ l of extract, 1 μ l of 10x denaturing buffer was added and the sample was incubated at 100°C for 10 minutes. Then, for the PNGaseF digest, 2 μ l G7-buffer, 2 μ l 10% NP-40 and 2 μ l PNGaseF was added to the sample whereas for the endoH digest, 2 μ l reaction buffer and 1 μ l endoH was added. Sample volume was adjusted to 20 μ l with ddH₂O, after which the sample was incubated at 37°C for one hour and analyzed by western blotting.

Antigen purification by IMAC

The GP4-Tm and GFP:GP4-Tm antigens were purified by IMAC on an Äkta Explorer® purification system fitted with a His Trap® Fast Flow column (1ml, GE healthcare). Seed extracts were prepared in binding buffer (50 mM phosphate buffer pH 7.4, 0.3 M NaCl, 20 mM imidazole, 0.1% CHAPS supplemented with an EDTA-free proteinase inhibitor cocktail (1 tablet per 50 ml; Roche)) by adding frozen and pulverized seeds to cold buffer at a fixed w/v of 1 g seeds per 70 ml buffer. The extract was vortexed and centrifuged for 30 minutes at 4°C and 40,000 g to remove cell debris. The supernatant was removed, passed through a GF-prefilter (Sartorius) and a Millex®-HP filter unit (Millipore) and then flowed over the IMAC column. After washing the column with binding buffer, antigens were eluted with 0.5 M imidazole (in binding buffer). To remove the imidazole, the antigens were washed with PBS on Amicon® Ultra centrifugal filters (Millipore, 3kDa molecular weight cutoff). Antigen purity was determined by densitometric analysis of Coomassie-stained polyacrylamide gels using the aforementioned imaging system and the Image Lab 3.0 software (Biorad). Antigen concentration was derived from the sample optical density at 280 nm measured by a NanoDrop® ND-1000 spectrophotometer.

Antigen purification by protein-A affinity chromatography

To purify antigens carrying an Fc chain (mFc3, mFc2a or pFc) protein-A affinity chromatography was performed as described in (De Buck et al., 2013) with minor modifications. Seed extracts were prepared in binding buffer (50 mM phosphate buffer pH 7.4, 0.3 M NaCl, 0.1% CHAPS supplemented with an EDTA-free proteinase inhibitor cocktail (1 tablet per 50 ml; Roche)) by adding frozen and pulverized seeds to cold buffer at a fixed w/v of 1 g seeds per 70 ml buffer. The extract was vortexed and centrifuged for 30 minutes at 4°C and 40,000 g to remove cell debris. The supernatant was removed, passed through a GF-prefilter (Sartorius) and a Millex®-HP filter unit (Millipore) and then flowed over the protein-A column (MabSelect SuRe 1ml; GE Healthcare) fitted onto Äkta Explorer® purification system. After washing the column with binding buffer, antibodies were eluted with 1 M arginine buffer (pH 2.7) and immediately neutralized with 1 M TrisHCl (pH 9.0). To remove the arginine, the antigens were washed with PBS on Amicon® Ultra centrifugal filters (Millipore, 3kDa molecular weight cutoff). Antibody concentration was derived from the sample optical density at 280 nm measured by a NanoDrop® ND-1000 spectrophotometer.

Antigen purification by gel filtration

After an initial chromatography step, antigens to be used in ELISA and in the vaccination trials were further subjected to gel filtration to remove remaining impurities. To this goal, IMAC or protein-A elution fractions were passed over a Superdex-200 column (GE Healthcare) fitted onto an Äkta Explorer® purification system with lipopolysaccharide-free PBS as the mobile phase. Removal of impurities and/or degradation products was assayed by Coomassie staining of polyacrylamide gels and whenever necessary, pure, gel-filtrated antigen was concentrated using Amicon® Ultra centrifugal filters (Millipore, 3kDa molecular weight cutoff). The antigen concentration was derived from the sample optical density at 280 nm measured by a NanoDrop® ND-1000 spectrophotometer.

ELISA to determine antigen accumulation in the seed

The antigen content of the seed of high accumulating lines was analyzed by ELISA. Seed extracts were prepared as detailed above, dilution series were made in duplicate with PBS as coating buffer, loaded onto Maxisorp 96-well plates (Nunc, Sigma Aldrich) and coated overnight at 4°C. Wells were washed three times with 270 µl washing buffer (PBS, 0.1% (v/v) Tween-20) and this was repeated between all subsequent steps. Remaining protein-free binding sites were blocked with 270 µl blocking buffer (PBS, 1% (w/v) bovine serum albumin (BSA, Sigma Aldrich), 0.1% (v/v) Tween-20) for

60 minutes at room temperature. 100 μ l primary antibody solution was added and plates incubated for 60 minutes at room temperature. Secondary antibody solution was added (100 μ l) and plates again incubated for 60 minutes at room temperature. Plates were then washed five times with washing buffer and 100 μ l 3, 3', 5, 5'-Tetramethylbenzidine Liquid Substrate (TMB, Sigma Aldrich) was added. After 10 minutes, color development was terminated by addition of 100 μ l 1 M HCl and plates were read at 450 nm in a VERSAmax tunable microplate reader. Standard curves were derived from duplicate two-fold serial dilutions of purified antigen, measured values were plotted against standard curves and antigen content of the seed calculated after correcting for antigen molecular weight. For analysis of the GP3-Tm fusion antigens, the mAb α -GP3 VII2D was used (diluted 25:1000 in blocking buffer) as primary antibody together with α -mouse IgG1 (gamma-chain)-HRP (from rabbit, Sigma Aldrich), diluted 1:5000 in blocking buffer, as the secondary antibody and protein-A purified GP3-Tm:pFc as standard. The GP4, GP4-Tm and sGP4-Tm containing antigens were analyzed with the α -GP4 mAb XVIII5G (diluted 25:1000 in blocking buffer) as primary antibody, the α -mouse IgG1 (gamma-chain)-HRP (from rabbit, Sigma Aldrich), diluted 1:5000 in blocking buffer, as the secondary antibody and protein-A purified GP4-Tm:pFc as standard. For all antigens carrying GP5 or GP5-Tm, a murine anti His₆-tag (Serotec) primary antibody was used (diluted 1:1000 in blocking buffer), together with an α -mouse IgG HRP-linked whole antibody (from sheep, GE healthcare) as secondary antibody (diluted 1:5000) and protein-A purified GP5-Tm:pFc as standard.

Immunization of mice

Six- to eight-week-old female C57BL/6 mice were obtained from Janvier (Belgium). Mice were housed under specific pathogen-free conditions. All animal experiments were approved by the Local Ethical Committee of Ghent University. Mice were vaccinated twice with a 3-week interval by subcutaneous injection with either 5 μ g purified GFP:GP4-Tm, GP4-Tm:mFc2a or sGP4-Tm:mFc2a mixed in a 1:1 ratio with the squalene-based oil-in-water adjuvant AddaVax (Invivogen). The antigens were isolated and purified from crude seed extracts by IMAC (GFP:GP4-Tm) or protein-A affinity chromatography (sGP4-Tm:mFc2a and GP4-Tm:mFc2a), followed by a gel filtration step as described before. During the trial, blood samples were collected at week zero (preimmune), two, four and six post immunization. Serum samples for IPMA and detection of virus neutralizing antibodies were incubated 30 minutes at 56 °C prior to freezing. Mice were euthanized after the final bleeding at week six.

Characterization of the antibody response in mice

For detection of the α -GP4 antibodies, ELISA was performed on serum samples with purified GP4-Tm:pFc as antigen. Maxisorp 96-well plates (Nunc, Sigma Aldrich) were coated overnight at 4°C with 1 μ g antigen per well (in 100 μ l PBS). Wells were washed three times with 270 μ l washing buffer (PBS, 0.1% v/v Tween-20) and this step was repeated between all following steps unless stated differently. Wells were blocked with 270 μ l blocking buffer (PBS, 0.1% Tween-20, 1% BSA) for two hours at room temperature. 100 μ l of four-fold serial dilutions (diluted in blocking buffer with an initial dilution of 1:50) of sera were added and plates incubated for 90 minutes at room temperature. Subsequently 100 μ l of goat anti-mouse IgG-HRP (Southern Biotech) diluted 1:10.000 in blocking buffer was added and plates incubated 90 minutes at room temperature. Wells were washed five times with 270 μ l blocking buffer and 100 μ l TMB (Sigma Aldrich) was added. After 10 minutes color development was terminated by addition of 100 μ l 1M HCl and plates were read at 450 nm in a VERSAmax tunable microplate reader.

To determine if the antibodies elicited by vaccination were able to recognize the PRRSV, the collected sera were tested via an IPMA as described elsewhere (Labarque et al., 2000). Virus neutralizing antibody titers were detected by a single-replication virus-neutralization test as described in (Vanhee et al., 2009) with minor modifications. Two-fold serial dilutions of sera (starting from a 1:10 dilution) were mixed with equal volumes of porcine alveolar macrophage-grown LV resulting in a final titer of 10^5 TCID₅₀/ml. Virus-sera mixtures were incubated for one hour at 37°C and transferred to a 96-well plate (50 μ l/well) with Marc-145 cells that were cultivated 48 hours prior to use. The inoculum was removed after one hour and replaced by medium, after which the cells were incubated for another 11 hours, fixed by drying and stored at -20°C. The cells were stained for PRRSV infection with mAb 13E2 (Van Breedam et al., 2011) against the nucleocapsid protein of PRRSV and peroxidase conjugated goat α -mouse polyclonal antibodies (Dako), followed by development with 3-amino-9-ethylcarbazole. The number of infected cells in each well was counted in three fields at 200 \times magnification, and expressed relative (%) to the mean number of infected cells for all mock conditions within the same experiment.

Immunization of piglets

Three piglets four weeks of age were purchased from a PRRSV-negative farm, their PRRSV-seronegative status was confirmed by IPMA (Labarque et al., 2000) upon arrival and animals were housed in isolation units with HEPA-filtered air. Heat lamps were provided during the first few weeks and piglets were fed grains sprinkled with milk powder and had access to water ad libitum.

The animal experiment was approved by the local ethical committee of the Faculty of Veterinary Medicine, Ghent University. The three piglets were injected twice in the neck muscles behind the ear (intramuscular) with an experimental vaccine following a prime – boost protocol. The vaccine consisted of an antigen cocktail containing 100 µg purified GP3-Tm:pFc, 100 µg purified GP4-Tm:pFc and 100 µg purified GP5-Tm:pFc (300 µg antigen in total) mixed with an oil-in-water adjuvant in a 1:1 ratio (Suvaxyn, Fort Dodge Animal Health). The antigens were isolated and purified from crude seed extracts by protein-A affinity chromatography followed by a gel filtration step as described before. The booster vaccination was administered at week four and blood samples (5 ml taken from the jugular vein with disposable syringes) were collected at a weekly basis from week zero (preimmune) onward. Serum samples for IPMA and detection of virus neutralizing antibodies were incubated 30 minutes at 56 °C prior to freezing. Piglets were euthanized at the end of the experiment by slow injection of an overdose (15 ml) of a central nervous system depressant (Natrium Pentobarbital 20%) in the jugular vein.

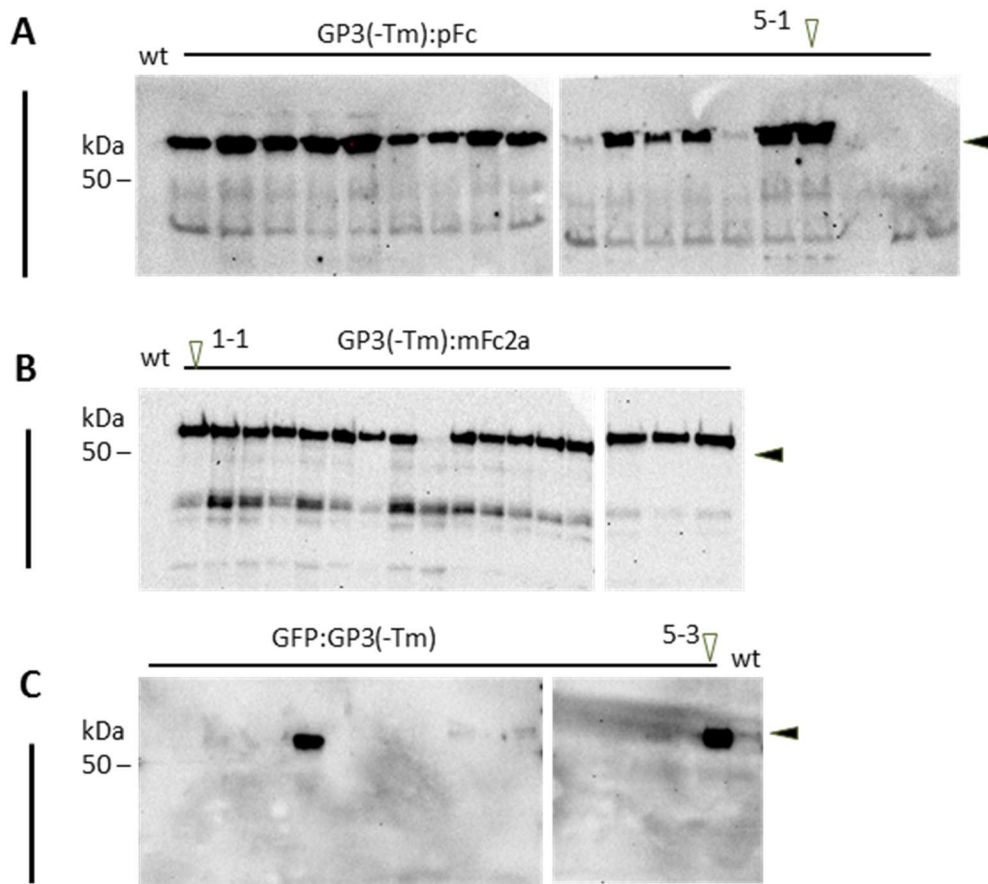
Characterization of the antibody response in piglets

Different ELISA protocols were used to measure the α -GP3, α -GP4 and α -GP5 antibodies in the piglet sera. To detect the α -GP3 and α -GP4 antibodies, purified GP3-Tm:mFc2a or GP4-Tm:mFc2a was used as antigen respectively. ELISA was further performed as described above except that 100 µl of polyclonal α -porcine IgG-HRP diluted 1:40.000 in blocking buffer was added to the wells to detect the α -GP3 or α -GP4 antibodies. To measure α -GP5 antibodies, the wells were coated with a seed extract containing the GP5-Tm. To this goal one gram of homozygous GP5-Tm seed stock was extracted with 10 ml of PBS, the extract diluted 1:100 in PBS and 100 µl containing 0.3 µg GP5-Tm added to the wells. Plates were incubated overnight at 4°C and protocol was followed as described for the α -GP3 and α -GP4 antibodies. To determine if the antibodies elicited by vaccination were able to recognize the PRRSV, the collected sera were tested via an IPMA as described elsewhere (Labarque et al., 2000). Virus neutralizing antibody titers were detected by a single-replication virus-neutralization test as described above.

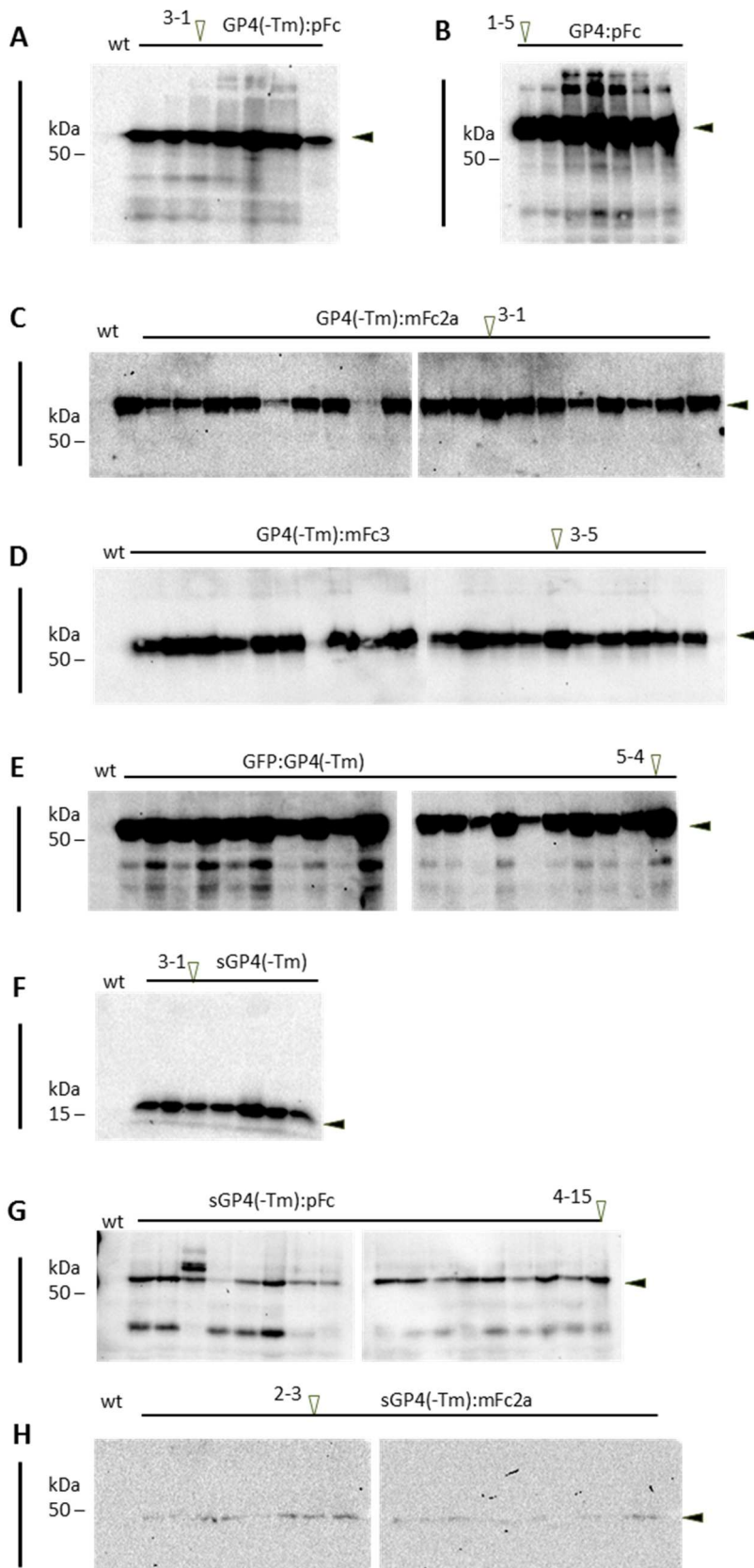
Acknowledgments

The authors thank Jonah Nolf, Els Van Lerberge and Janne Six for practical assistance with the experiments, the research groups of J. Grooten and H. Nauwynck for assistance with the vaccination trials and Annick Bleys for help in preparing the manuscript.

Supplemental information

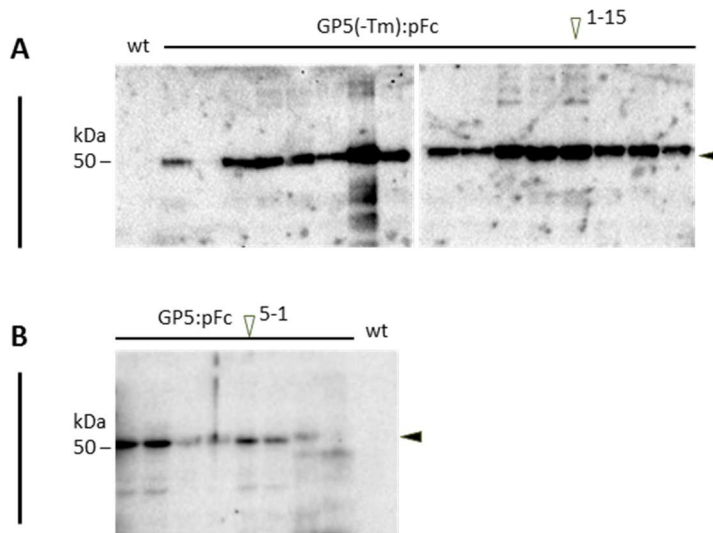


Supplemental figure S1. T2-analysis and inter-transformant variation in accumulation level of the GP3 derived antigens. For each antigen, 20 transformed T2 seed stocks were analyzed by western blot to determine their antigen content, except for the GP3(-Tm):mFc2a, for which only 17 T2 seed stocks were screened. After visual inspection of the western blots, seven stocks were selected, analyzed for single locus T-DNA integration and the stock with the highest antigen content and the T-DNA integrated at a single locus was retained, selfed and further propagated. A hollow arrow (and their corresponding number) marks the propagated stocks. Seed extracts were prepared at a fixed w/v ratio, subjected to western blotting and analyzed with (A) α -pig IgG-HRP to detect the GP3(-Tm):pFc, (B) an α -mouse IgG-HRP conjugate to visualize the GP3(-Tm):mFc2a and (C) an antibody against the His-tag to probe for the GFP:GP3(-Tm). 10 μ l was loaded and gels were run under reducing conditions. The position of the antigens on the blots is indicated with a black arrow. Abbreviations used: kDa, kilodalton; wt, wild-type seed extract.



Supplemental figure S2. T2-analysis and inter-transformant variation in accumulation level of the GP4 derived antigens. The GP4(-Tm) and GP4 transgenic lines were propagated to homozygosity by Dr. A. De Paepe and provided as such. Information on inter-transformant variation was not at hand for these lines. For each of the other antigens, 20 transformed T2 seed stocks were analyzed by western blot to determine their antigen content, except for the sGP4(-Tm):pFc, for which only 17 T2 seed stocks were screened and the GFP:GP4 that was screened by ELISA (data not shown). After visual inspection of the western blots, seven stocks were selected, analyzed for single locus T-DNA integration and the stock with the highest antigen content and the T-DNA integrated at a single locus was retained, selfed and further propagated. A hollow arrow (and their corresponding number) marks the propagated stocks. Seed extracts were prepared at a fixed w/v ratio, 10 μ l was loaded onto the gels that were run under reducing conditions. The seven GP4(-Tm):pFc (A) and GP4:pFc stocks (B) selected for single locus determination and detected with the α -GP4 mAb XVI 11C (Costers et al., 2010). (C) The twenty screened GP4(-Tm):mFc2a stocks visualized with α -mouse IgG-HRP. (D) The twenty screened GP4(-Tm):mFc3 stocks assayed with goat α -mouse IgG3-HRP. (E) Twenty screened GFP:GP4(-Tm) stocks probed with an antibody against the His₆-tag. (F) The seven sGP4(-Tm)

stocks selected for single locus determination and detected with the α -GP4 mAb XVI 11C (Costers et al., 2010). (G) Twenty screened sGP4(-Tm):pFc stocks probed with an antibody against the His₆-tag. (H) Twenty screened sGP4(-Tm):mFc2a stocks visualized with goat α -mouse IgG-HRP. Abbreviations used; kDa, kilodalton; wt, wild-type seed extract.



Supplemental figure S3. T2-analysis and inter-transformant variation in accumulation level of the GP5 derived antigens. The GP5(-Tm) and GP5 transgenic lines were propagated to homozygosity by Dr. A. De Paepe and provided as such. Information on inter-transformant variation was not at hand for these lines. For the GP5:pFc and GP5(-Tm):pFc twenty transformed T2 seed stocks were analyzed by western blot to determine their antigen content. After visual inspection of the western blots, seven stocks were selected, analyzed for single locus T-DNA integration and the stock with the highest antigen content and the T-DNA integrated at a single locus was retained, selfed and further propagated. A hollow arrow (and their corresponding number) marks the propagated stocks. Seed extracts were prepared at a fixed w/v ratio, 10 μ l was loaded onto the gels that were run under reducing conditions and detected with an antibody against the His₆-tag. (A) The twenty screened GP5(-Tm):pFc stocks. (B) The seven GP5:pFc stocks selected for single locus determination.

Supplemental table 1. Antigen accumulation level

Antigen	Line	mg/g DS(+/- SD)	% of TSP
GP3-Tm:pFc	Ho 5-1-9	2.905 +/- 0.289	0.96 +/- 0.09
GP3-Tm:mFc2a	Ho 1-1-8	6.185 +/- 2.235	1.81 +/- 0.36
GFP:GP3-Tm	Ho 5-3-2	7.858 +/- 0.404	2.74 +/- 0.16
GP4	Ho 5/1-5	0.245 +/- 0.033	0.08 +/- 0.01
GP4:pFc	Ho 1-5-3	0.485 +/- 0.048	0.14 +/- 0.01
GFP:GP4	Se 1-2	0.467 +/- 0.038	0.15 +/- 0.02
GP4-Tm	Ho 5/3-4-1	0.170 +/- 0.015	0.05 +/- 0.01
GP4-Tm:mFc3	Ho 3-5-13	0.151 +/- 0.019	0.05 +/- 0.01
GP4-Tm:pFc	Ho 3-1-4	2.064 +/- 0.086	0.67 +/- 0.04
GP4-Tm:mFc2a	Ho 3-1-4	5.153 +/- 0.276	1.59 +/- 0.11
GFP:GP4-Tm	Ho 5-4-5	7.377 +/- 0.334	2.36 +/- 0.30
sGP4-Tm	Se 3-6	0.019 +/- 0.003	0.005 +/- 0.001
sGP4-Tm:pFc	Se 4-15	0.183 +/- 0.014	0.06 +/- 0.002
sGP4-Tm:mFc2a	Ho 2-3-1-1	0.133 +/- 0.011	0.04 +/- 0.01
GP5	Ho 3/2-5-1	<i>0.273</i> +/- -	<i>0.10</i> +/- -
GP5:pFc	Ho 5-1-3-2	<i>0.015</i> +/- -	<i>0.01</i> +/- -
GP5-Tm	Ho 4/9-E-7	0.258 +/- 0.059	0.09 +/- 0.005
GP5-Tm:pFc	Se 1-15	2.333 +/- 0.077	0.83 +/- 0.022

Table S1. For each antigen given values are the mean of three biochemical repeats. Abbreviations used: Ho, homozygous seed stock; Se, segregating seed stock; DS, dry seed; SD, standard deviation; TSP, total soluble protein. Numbers in *italic* are derived from signal quantification after western blotting via the His₆-tag, since the GP5 and GP5:pFc antigen could not be detected by ELISA.

Supplemental table 2. Extension PCR data

Project	Target of amplification	PCR reaction	Primer		T ann	
			Forward	Reverse	5x	20x
Cloning GP3(-Tm):pFc	GP3(-Tm)		1 ropir1	ropir12	65	72
			2 ropir1	ropir4	64	69
			3 ropir1	ropir5	61	72
Cloning GP4(-Tm):pFc	GP4(-Tm)		1 ropir9	ropir10	58	62
			2 ropir1	ropir4	62	69
			3 ropir1	ropir5	60	69
Cloning sGP4(-Tm):pFc	sGP4(-Tm)		1 ropir1	ropir2	54	71
			2 ropir1	ropir5	60	69
Cloning GP4:pFc	GP4		1 ropir9	ropir10	58	62
			2 ropir1	ropir4	62	69
			3 ropir1	ropir5	60	69
Cloning GP5(-Tm):pFc	GP5(-Tm)		1 ropir9	ropir10	58	62
			2 ropir1	ropir4	62	69
			3 ropir1	ropir5	60	69
Cloning GP5:pFc	GP5		1 ropir9	ropir10	58	62
			2 ropir1	ropir4	62	69
			3 ropir1	ropir5	60	69
Cloning GP4(-Tm):mFc3	GP4(-Tm)		1 ropir1	mFc3R3	52	70
Cloning GP3(-Tm):mFc2a	GP3(-Tm)		1 ropir9	mFc2aR2	54	63
Cloning GP4(-Tm):mFc2a	GP4(-Tm)		1 ropir1	mFc3R3	52	70
Cloning sGP4(-Tm):mFc2a	sGP4(-Tm)		1 ropir9	mFc3R3	52	63
Cloning GFP:GP4(-Tm)	GFP:GP4(-Tm)		1 ropir1	JaR5	53	69
			2 ropir1	JaR6	65	69
Cloning GFP:GP3(-Tm)	GP3(-Tm)		1 JaF7	JaR8	58	71

Table S2. (Multistep) extension PCR data. Are given, the cloning project involved, the target of amplification, the forward and reverse primer used in each reaction, the number of consecutive extension PCRs performed for each project and the annealing temperature of the first five and last twenty cycles. The Phusion® High-Fidelity DNA Polymerase (Finnzymes) was used for all reactions according to the manufacturer's guidelines. Abbreviations used; T ann, annealing temperature.

Supplemental table 3. Primer sequences

Primer name	Sequence (5' to 3')
ropir1	CCGGAATTCCGGCCACCATGGCAAACAAGCTTTTCCT
ropir2	AGATGGGTGTGGGGGGTTCGATGTCGCTTCCTCTTGGAACGAGGTGATGGTGATGGTGGTG
ropir4	AGCAAATTCGGGACAGATGGGTGTGGGGGGTTCGATGTCGCTTCCTCTTGGAACGAGGTGATGG
ropir5	TACGCCTAGGACCTCTGCAGCTGGGCATGAGCAAATTCGGGACAGATGGGTGT
ropir9	TGAATTCCCACCATGGCAAACAAGCTTTTCCT
ropir10	GCTTCCTCTTGGAACGAGGTGATGGTGATGGTGGTGGCTTGGTGTAAGGA
ropir12	GCTTCCTCTTGGAACGAGGTGATGGTGATGGTGGTGGTCCAAAGTGAACC
mFc2aR2	TGGACCCCGCGGTTCCAAGTGGAACCAAGTTGCC
mFc3R3	GTATCCGCGGGCTTGGTGTAAGGAAGTGGA
JaF7	ATCACTCCGCGGACGAAGTTGATTCAAACCTCCAGCTCTACGCTATGCTTT
JaR5	TGGCGAACAAGCAAGCAATGGTTGTAGCCACCTCATTGCGCTTGGTGTAAGGAAGTGGA
JaR6	GGTGGATCCTTAAAGCTCGTCTTTGATTGCCAAGAGAATGGCGAACAAGCAAGCAATG
JaR8	GGGTGGATCCTCAAAGCTCGTCTTTTCCAAGTGGAACCAAGTTGCCACCG

Chapter III

Glycoengineering in *Arabidopsis* seeds, an introduction.

Robin Piron, Ann Depicker and Nico Callewaert.

R.P. wrote the chapter, A.D. and N.C. edited it.

Glycosylation

Protein glycosylation, or the covalent linkage of an oligosaccharide side chain to the amino acid backbone of a protein, represents the most widespread PTM. One third of all therapeutic proteins approved for human use are glycosylated and it is estimated that this also holds true for 50 % of all human proteins (Campbell & Yarema, 2005; Saint-Jore-Dupas et al., 2007; Walsh & Jefferis, 2006). Glycosylation profoundly influences several critical characteristics of a protein such as solubility, stability, susceptibility to proteases, serum half-life, function and activity (Jacobs & Callewaert, 2009; Nagels et al., 2012b). Glycosylation is of such importance that the platform used for therapeutic protein production is often chosen based on its capacity to decorate the therapeutic with those glycans that are required for its optimal performance (Saint-Jore-Dupas et al., 2007). This is the main reason why mammalian cells are preferred over organisms such as yeasts for the production of therapeutic proteins. Indeed, the glycosylation performed by mammalian cells is very similar (but not identical) to that of humans (Jacobs & Callewaert, 2009).

There are, however, several drawbacks associated with protein production in mammalian cells. A glycoprotein that enters the ER will be decorated with a defined glycan structure at specific sites. Once folded, the glycoprotein will pass through the secretory pathway where the attached glycan structures will be modified by the sequential action of an array of glycosidases and glycosyltransferases. Differences in accessibility of the glycans on the protein combined with the non-template driven nature of the glycosylation machinery introduce a large heterogeneity in the final glycan structures. One glycoprotein will thus exist as several 'glycoforms', sharing the same protein backbone but decorated with different glycans (Jacobs & Callewaert, 2009). This heterogeneity is highly undesirable from both a practical as a regulatory point of view. Different glycoforms will show differential behavior during downstream processing and *in vivo*. This requires that each glycoform is extensively characterized, which complicates therapeutic protein production.

Cultivation of mammalian cells is a costly process, expensive high-tech equipment is needed and the cultivation process needs to be closely monitored to guarantee batch-to-batch reproducibility (Saint-Jore-Dupas et al., 2007). Moreover, mammalian cells are prone to contamination with pathogens such as viruses and prions and glycan heterogeneity is rather high (Bosch et al., 2013; Bosch & Schots, 2010). To avoid these drawbacks, alternative production platforms are being investigated such as, for example, plants. An advantage about plants in the context of therapeutic glycoprotein production is that they are not susceptible to mammalian pathogens, production is scalable and possibly cheap and last but not least, plant glycans are more homogeneous than mammalian glycans, with all the associated benefits (Faye et al., 2005; Gomord & Faye, 2004). Also, plants seem to be

quite tolerant to alterations of their N-glycosylation machinery (Bosch et al., 2013). This sparked the production of therapeutic proteins in several plant species and tissues, ranging from plant cell cultures to entire plants and this via stable and/or transient expression, which has been covered in more detail in the section ‘Transgenic plants as biofactories’.

Plant N-glycosylation

Proteins can be glycosylated by attachment of a glycan structure to an asparagine residue in the consensus sequence Asn-X-Ser/Thr, with “X” being any amino acid except for proline, and this is referred to as N-glycosylation. Additionally, glycans can be transferred to the hydroxyl groups of serine, threonine and for plants, also hydroxyproline. This type of glycosylation, termed O-glycosylation, will not be detailed here but is extensively reviewed by Gomord et al. (2010). In many aspects, plant O- and N-glycosylation are alike; both have a profound influence on the biochemical characteristics and biological function of proteins, they are considered to be different between humans and plants and they can be immunogenic. As O-glycosylation is not within the scope of the later presented work, we will here focus on N-glycosylation in plants and contrast this with the N-glycans found on human proteins.

The entire N-glycosylation machinery is located in the secretory pathway (Figure 1) and starts with the co-translational addition of the oligosaccharide precursor $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ onto the nascent protein at N-x-S/T attachment sites. Once transferred to the glycoprotein, the precursor recruits chaperones and oxidoreductases that aid the glycoprotein to reach its correct fold (Pearse & Hebert, 2010). During this process, the quality control mechanism in the ER trims the precursor down to $\text{Man}_8\text{GlcNAc}_2$ (Figure 1, glycan 1) on correctly folded glycoproteins and these are subsequently released to continue towards the Golgi apparatus (Pearse & Hebert, 2010). These initial steps are shared between all higher eukaryotes; it is only after transit through the Golgi that species-specific N-glycans arise and these are known as ‘complex N-glycans’. When a glycoprotein reaches the *cis*-Golgi the mannose residues on its glycans are further trimmed and an N-acetylglucosamine residue (GlcNAc) is added to each branch. The resulting $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ (Figure 1, glycan 5) is basically the branching point from where plant and human N-glycosylation differ (Bosch et al., 2013).

In plants, further passage through the *medial* and *trans*-Golgi adds a bisecting β 1,2-xylose to the core mannose and an α 1,3-fucose to the innermost GlcNAc, respectively (Figure 1). These modifications (one or both) are omnipresent on complex plant N-glycans, but are absent in humans and are potentially problematic (see later). Most recombinant glycoproteins that are produced in plants will

thus predominantly carry GnGnXF N-glycans⁴ (Figure 1, glycan 7; (Bosch et al., 2013)). Nonetheless, two additional modifications can be found on secretory glycoproteins. Proteins that are targeted to the vacuole tend to carry MMXF glycans (Figure 1, glycan 10) and it is believed that they are the product of vacuolar hexosaminidases that trim off terminal GlcNAc residues (Liebminger et al., 2011). The resulting glycan structures are known as ‘paucimannosidic N-glycans’ and are considered to be a hallmark of proteins with a vacuolar localization. These paucimannosidic N-glycans are also present on some secreted glycoproteins and recently, a hexosaminidase associated with the plasma membrane was discovered (Liebminger et al., 2011). Alternatively, although this generally occurs at very low level on recombinant glycoproteins, the terminal GlcNAc residues of some secretory proteins are further extended with both an α 1,4-fucose and a β 1,3-galactose (Fitchette-Laine et al., 1997). This structure, known as the ‘Lewis A epitope’ (Figure 1, glycan 9) can also be found on some human glycans (Sourrouille et al., 2008).

In humans, integration of additional GlcNAc residues in the Golgi can lead to further branching such that the final complex N-glycan can be anywhere from bi- to tetra-antennary (Figure 2B). Another major difference is that terminal GlcNAc residues can be further extended with galactose moieties which are then capped with a sialic acid. Besides sialic acid, also a terminal ‘Lewis X epitope’ is sometimes present. Substituting the plant β 1,2-xylose, a bisecting β 1,4-GlcNAc can be found attached to the core mannose instead, and the innermost GlcNAc often carries a fucose in an α 1,6 linkage as opposed to the α 1,3-fucose encountered in plants.

To summarize, plants are capable of producing glycoproteins decorated with complex, bi-antennary N-glycans that are mainly GlcNAc-terminated and rather homogeneous when compared to their mammalian counterparts (compare Figure 2A and 2B). However, these glycans carry plant specific fucose and xylose residues and lack some typical mammalian residues. Furthermore, these characteristics seem to be conserved across the plant-kingdom (Bosch & Schots, 2010; Nagels et al., 2011a).

⁴ Bi-antennary plant N-glycans are abbreviated in the tekst using the nomenclature proposed on www.proglycan.org

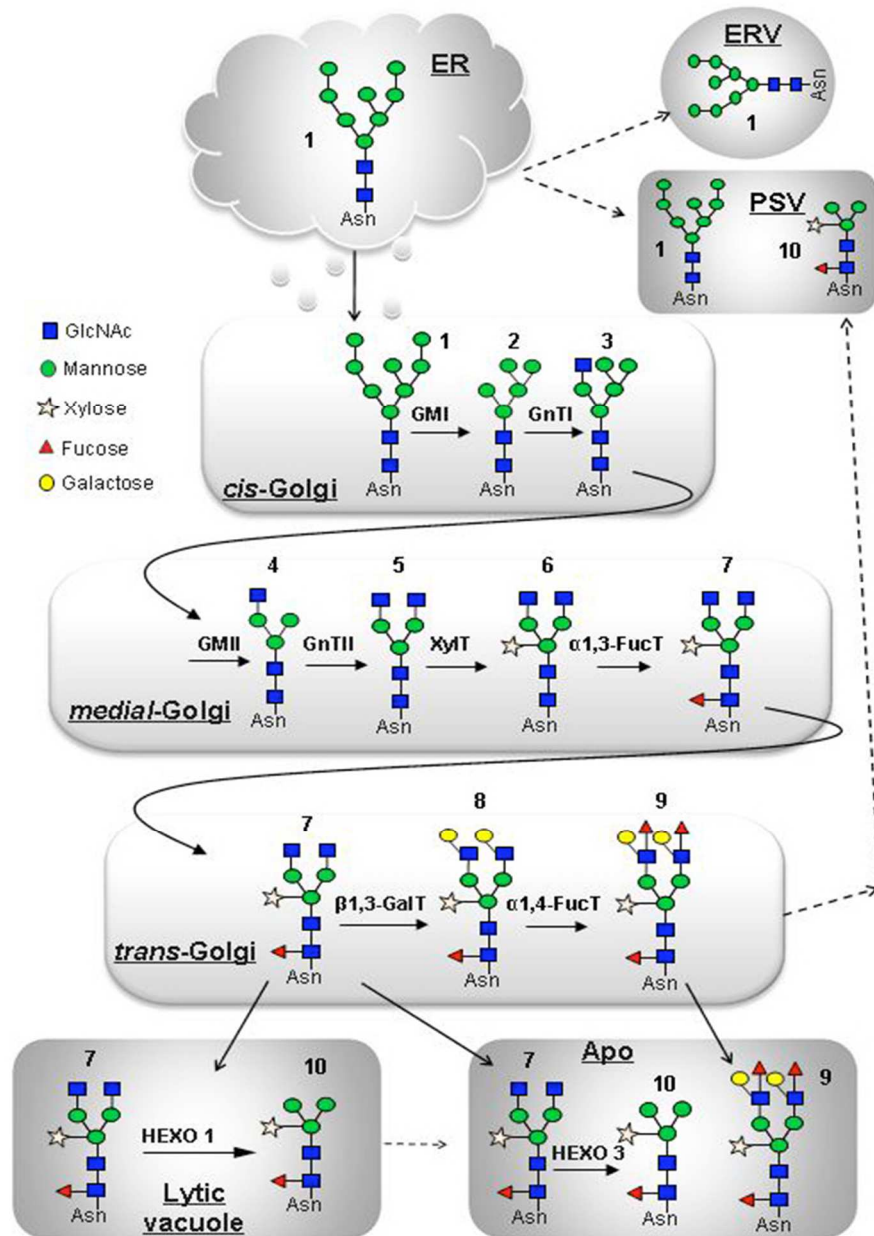


Figure 1. The plant N-glycosylation pathway. Figure taken from Bosch et al. (2013). A correctly folded glycoprotein will exit the ER carrying Man₈GlcNAc₂ glycans (1) and these are subsequently modified by the sequential action of glycosidases and glycosyltransferases located in the different compartments of the secretory pathway. Paucimannosidic glycans are a hallmark of vacuolar proteins but can also be found in minor quantities on secreted proteins. Hexosaminidase processing is a likely mechanism for their formation. Abbreviations used: ER, endoplasmic reticulum; ERV, ER-derived vesicle; PSV, protein storage vacuole; Apo, apoplast; GMI, Golgi- α -mannosidase I; GnTI and II, N-acetylglucosaminyltransferase I and II respectively; GMII, Golgi- α -mannosidase II; XylT; β 1,2-xylosyltransferase; α 1,3-FucT, core α 1,3-fucosyltransferase; β 1,3-GalT, β 1,3-galactosyltransferase; α 1,4-FucT, α 1,4-fucosyltransferase; HEXO1 and HEXO3, β -N-acetylhexosaminidase 1 and 3 respectively. All glycan structures are depicted according to the recommendations of the international consortium of functional glycomics⁵.

⁵ www.functionalglycomics.org

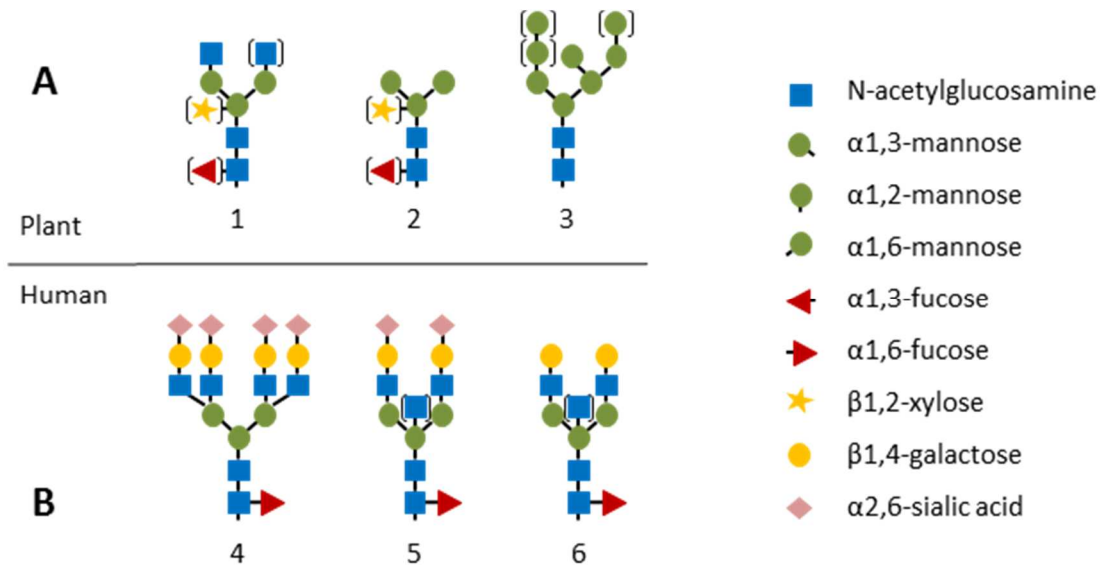


Figure 2. Overview of the difference between plant and human complex N-glycans. (A) Typical glycan structures found on therapeutic proteins produced in plants. Fucose and/or xylose bearing complex glycans with one or two terminal GlcNAcs (1) are the predominant structures present on secretory recombinant proteins expressed *in planta*. Paucimannosidic structures with fucose and/or xylose but lacking terminal GlcNAcs (2) are sometimes also found, together with small quantities of high-mannose glycans that can have from five to eight mannoses but never carry fucose and xylose (3). **(B)** Typical glycan structures found on human secretory proteins. Anywhere from two (5) to four antennae (4) can be present on a human glycans and these tend to carry terminal sialic acids. An exception are the glycans of human IgG Fc domains that are predominantly biantennary with terminal galactoses (6) and only a minor fraction of these glycans is further extended with sialic acid (5). Fucose is always attached through an α 1,6-linkage to the proximal GlcNAc and instead of β 1,2-xylose, a bisecting GlcNAc can be present on bi-antennary, but not tri- or tetra-antennary structures. All glycan structures are depicted according to the recommendations of the international consortium of functional glycomics.

Immunogenicity of plant N-glycans

Plant N-glycosylation in the context of therapeutic protein production is a controversial topic, mainly because of the two sugar residues, β 1,2-xylose and α 1,3-fucose, that are ubiquitously present on complex plant N-glycans, but are absent in human glycans and thus can be immunogenic (Figure 2). There are several observations to support their immunogenicity, namely that vaccination of lab animals with plant glycoproteins raises antibodies specific for these glycans and that a whole range of plant-glycan specific antibodies, covering almost all immunoglobulin isotypes, can be found in the blood of humans. These observations are further detailed in the following paragraphs.

Firstly, immunization with proteins decorated with complex plant N-glycans consistently generates antibodies specific for both β 1,2-xylose and α 1,3-fucose in goats, rabbits and rats, but not mice who, depending on their genetic background (e.g. BALBc mice), do not seem to respond against these residues (Bardor et al., 2003; Chargelegue et al., 2000; Jin et al., 2008a; Kurosaka et al., 1991). Secondly, anti α 1,3-fucose and β 1,2-xylose IgG and IgM antibodies are present in the blood of 25 and 50% of non-allergic human blood donors respectively (Bardor et al., 2003). Thirdly, β 1,2-xylose and α 1,3-fucose constitute carbohydrate epitopes that can be recognized by IgE antibodies and both are present on a wide range of allergens, hence they are termed 'cross-reactive carbohydrate determinants' (CCDs; (Aalberse & van Ree, 1996)).

Apparently, IgE antibodies specific for plant N-glycans are quite widespread with a high prevalence among the pollen-allergic and it is believed that they arise in response to insect stings and/or exposure to pollen (van Ree, 2002). More than 20% of persons with an allergic respiratory disease were shown to carry plant-specific CCD-IgE antibodies and this number rose to about 70% for persons with multiple pollen allergies (Mari, 2002). Strikingly, plant-specific CCD-IgE could also be found in 5% of the non-allergic persons. CCD-IgE could potentially be very dangerous as cross-linking of receptors by IgE molecules on effector cells such as mast cells causes mediator release that can result in an allergic reaction. It should be noted however that presence of an IgE showing biological activity (mediator release) does not necessarily translate into a severe allergic reaction and this apparent contradiction is reflected in the ongoing debate about the clinical relevance (role in causing allergic symptoms) of these plant-specific CCD-IgE antibodies (van Ree, 2002).

Plant-specific CCD-IgE antibodies from the sera of allergic persons are capable of causing histamine release from basophils *in vitro* and although this is only true for a subset of the tested sera, this has been regarded as evidence of their biological activity *in vivo* (Bublin et al., 2003; Foetisch et al., 2003). However, an argument against the clinical relevance of CCD-specific IgE is that topical application of plant glycoproteins carrying complex N-glycans is generally well tolerated, even in the

allergic (Ma et al., 1998; Mari et al., 2008; Zeitlin et al., 1998). For example, when allergic patients were screened for reactivity against a rice-produced human lactoferrin carrying complex plant N-glycans, poor or no biological activity of the lactoferrin-binding IgE antibodies was found. Although almost all the sera with plant-specific CCD-IgE antibodies could bind to the recombinant lactoferrin, only half of the sera tested *in vitro* were found to be biologically active releasing histamine and when compared to a major grass pollen allergen, the concentrations required for biological activity were 5 to 6 orders of magnitude higher. Eventually patients that tested positive for biological activity *in vitro* were also submitted to a skin prick test and an oral challenge with the lactoferrin, but without any adverse reactions. This led the authors to conclude that the plant-specific CCD-IgE antibodies had poor or no biological activity and lacked any clinical relevance whatsoever (Mari et al., 2008).

Several explanations can be given for this apparent lack of clinical relevance. A prerequisite for biological activity of the IgE is that it binds with sufficient affinity to its target and that this is at least bivalent so that crosslinking occurs on the surface of the effector cell and mediator is released (Altmann, 2007; Foetisch et al., 2003). However, it was recently reported that human IgE and IgG specific for plant-CCDs have high affinities for their carbohydrate epitopes, debasing this affinity argument (Jin et al., 2008b). This led Altmann (2007) to doubt the CCD-valency of the recombinant lactoferrin used during the study performed by Mari et al. (2008). Nevertheless detailed glycan analysis showed that two out of three N-glycan sites were occupied with MMXF and MUXF, representing 83 % of the total glycan pool of which no less than 93 % carried xylose or fucose (Fujiyama et al., 2004). Also, only about 21 % of patients carrying plant-specific CCD-IgE tested positive in a skin-prick test when horse radish peroxidase was applied and these generally presented very mild symptoms (Mari, 2002). Horse radish peroxidase carries six identical MMXF N-glycans (Kurosaka et al., 1991). Given these results it seems reasonable to discard valency as an explanation for low plant-glycan CCD-IgE activity. What is more probable is that an initial insect sting and/or pollen exposure generates CCD-IgE, but humans are desensitized by the continuous contact with and ingestion of plant proteins decorated with CCDs. In this respect, the high prevalence of plant-glycan specific IgG and IgM in the blood of humans (Bardor et al., 2003) offers a plausible explanation for this desensitization. The low titers of plant-glycan specific IgG and IgM, present in the blood because of continuous exposure to plant glycoproteins, could compete for binding to complex plant glycans. This would block interaction of CCD-IgE with its target thus reducing its biological activity and concomitantly, its clinical relevance (Altmann, 2007).

To summarize, plant-specific CCD-IgE is quite widespread in the human population but seems of little clinical relevance for the topical or oral application of a PMP carrying complex N-glycans.

Nonetheless, this picture changes when parenteral delivery of a PMP is envisaged, especially when considering a therapeutic protein that needs to be administered repeatedly. One could argue that many if not all persons with plant-specific CCD-IgE also have plant-specific CCD-IgG in their blood stream that block the biological activity of the CCD-IgE and that plant-CCDs are thus rather innocuous. However a recent incident serves to remind us that one should not thread lightly with respect to CCDs. Severe hypersensitivity reactions were reported for cetuximab, a recombinant monoclonal antibody approved for cancer therapy and delivered via the parenteral route. Most subjects who had a hypersensitivity reaction were shown to have CCD-IgE antibodies capable of binding to cetuximab in their serum before therapy. These antibodies were specific for galactose- α 1,3-galactose, a CCD that is present on glycoproteins from most non-primate mammals and was introduced on cetuximab as an artefact of its production in the mouse myeloma cell line SP2/0 (Chung et al., 2008).

The previous discussion clarifies that, in the context of therapeutic protein production for parenteral delivery, plant-specific sugar residues should be avoided whenever possible. The widespread presence of antibodies specific for plant glycans in the serum of patients might result in an accelerated clearance of the therapeutic and, in a worst-case scenario, might provoke a severe hypersensitivity reaction.

‘Humanizing’ the plant N-glycosylation pathway

Major efforts have been done to ‘humanize’ plant N-glycans, more specifically by avoiding the addition of immunogenic glycans and/or introducing human glycan residues into the plant N-glycosylation pathway. Several strategies have been pursued to obtain these goals. The N-glycosylation pathway only differs from the Golgi apparatus onward, so the most straightforward manner to avoid immunogenic plant glycans is to retain the therapeutic in the ER. This is in theory easily achieved by adding the ER-retention sequence ‘H/KDEL’ to the C-terminus of the protein (Gomord et al., 1997b). One should be aware that the KDEL sequence is actually a retrieval sequence so the therapeutic can cycle between the ER and the *cis*-Golgi where it might come into contact with mannosidases and glycosyltransferases. Hence if it is correctly retrieved it will present glycoforms that can vary from Man₉GlcNAc₂ to Man₅GlcNAc₂ with only minor levels of Golgi typical modifications. These are collectively termed ‘high-mannose’ N-glycans⁶ and are typical for ER-localized proteins. Overexpression and retention of a recombinant glycoprotein in the ER can give

⁶ These will further be depicted in the text as ‘Man_x’, with ‘x’ being the number of mannose residues

rise to ER-derived vesicles and glycoproteins present in these vesicles also carry typical high-mannose N-glycans (Loos et al., 2011b). However, the functionality of ER-retention tags is not absolute and some proteins can escape the retrieval mechanism, which results in increased glycan heterogeneity and, worse, glycoforms carrying immunogenic complex N-glycans (He et al., 2012b). Also, for some tagged proteins, complex N-glycans represent the majority of the glycoforms indicating that the retention tag is non-functional and this supports the consensus that retention in the ER by a retrieval sequence is protein specific (Gomord et al., 2010).

Another strategy to avoid the incorporation of β 1,2-xylose and α 1,3-fucose is to inactivate the enzymes responsible for their transfer. This has been achieved by two methods. For *Arabidopsis thaliana* and the moss *Physcomitrella patens*, mutants were obtained defective in the genes coding for xylosyl and fucosyltransferase, respectively by screening mutant libraries and by targeted knockout via homologous recombination (Huether et al., 2005; Koprivova et al., 2004; Strasser et al., 2004). Homologous recombination is exceptionally straightforward in the moss *Physcomitrella*. However for other species these resources were not available and the fact that plant genes tend to be found in multigene families further complicated matters. Hence, gene silencing was employed to prevent the production of active enzymes and this method has proven more or less successful for a range of plant species including *Nicotiana benthamiana*, *Medicago sativa* and *Lemna minor* (Cox et al., 2006; Sourrouille et al., 2008; Strasser et al., 2008) and a number of the different lines lacking fucose and xylose obtained by these methods were used for further manipulations of the plant N-glycosylation pathway.

For example, a silenced *N. benthamiana* line was combined with human β 1,4-galactosyltransferase to produce the anti-HIV 2G12 monoclonal antibody devoid of xylose and fucose but with terminally galactosylated N-glycans, a prerequisite for the subsequent addition of terminal sialic acids (Strasser et al., 2009). The galactosyltransferase had been expressed *in planta* before, but only mixtures of galactosylated complex glycans still bearing some fucose and xylose were obtained, highlighting the importance of the fucosylation and xylosylation deficient lines (Bakker et al., 2001; Bakker et al., 2006; Palacpac et al., 1999).

By now, most of the structures typical for human N-glycans have been introduced *in planta* (Figure 3). Often by tedious coordinated multigene expression as exemplified by the work of Castilho et al. (2010) who managed to produce a sialylated version of the 2G12 anti-HIV monoclonal antibody in silenced *N. benthamiana* plants. Genes coding for the biosynthesis, activation, transport and transfer of sialic acid onto the N-glycan were transiently expressed in a coordinated manner resulting in no less than six proteins that were targeted to different subcellular locations. Another major

achievement was the introduction of the transferases responsible for the addition of branching GlcNAcs in silenced *N. benthamiana* lines. Both stable transformation and transient expression of the two genes coding for the transferases resulted in the formation of bi-, tri- and tetra-antennary N-glycans on the reporter protein erythropoietin C-terminally fused to an IgG Fc domain (Castilho et al., 2011; Nagels et al., 2012a; Nagels et al., 2011b). However, it should be noted that in all cases the introduced conversions were far from completely efficient, and more complex glycan mixtures were formed than encountered in the wild type plants.

Besides sialylation and multi-antennary structures, other human N-glycan structures have been introduced *in planta*, such as bisecting GlcNAc (Castilho et al., 2011; Rouwendal et al., 2007), core fucose in a α 1,6-linkage (Forthal et al., 2010) and the Lewis X epitope, which differs from the Lewis A epitope by the linkages through which the galactose and fucose are attached to the terminal GlcNAc (Lewis A: Fuc- α 1,4(Gal- β 1,3)GlcNAc, Lewis X: Fuc- α 1,3(Gal- β 1,4)GlcNAc; (Rouwendal et al., 2009)). The only remaining feature of human N-glycans that has not been introduced in plants yet, are multi-antennary, terminally sialylated N-glycans (Nagels et al., 2011b). As both components that make up this structure (multiple antennae and sialylation) were individually successfully integrated in the plant N-glycosylation machinery, it is plausible to assume that this is only a matter of time until both elements are combined. Although theoretically feasible, this once more presents a practically daunting task requiring the balanced expression of many genes whose products will need to be in the correct subcellular compartment at the right moment and in the needed quantity. Moreover, the hexosaminidase problem has not been satisfactorily solved, and this would require several more gene knock-outs.

The previous paragraphs reflect the research effort that has been done up to today to ‘humanize’ plant glycans and although major advances have been made during the years, some unresolved issues still remain. Firstly, most modification are not entirely effective and this can actually augment glycan heterogeneity when compared to the wild type plant, which is highly undesirable from both a regulatory as a commercial point of view. Secondly, some topics, e.g. the hexosaminidases have not been touched yet, and doing so will only further enhance the complexity of the final ‘humanized’ plants. Thirdly, although proof-of-concept has been delivered for most of these modifications by co-expression of reporterproteins, this has always been done on a lab scale and it will be extremely difficult to translate these results into industrial applications. Taken into account that not all commercially interesting crops are easily transformed, together with long regeneration times, polyploidy and the need for multiple crosses to integrate all favorable attributes into one elite cultivar, development times of ten years plus, are not exaggerated and actually, most efforts by industry to achieve this have been halted due to not economically feasible.

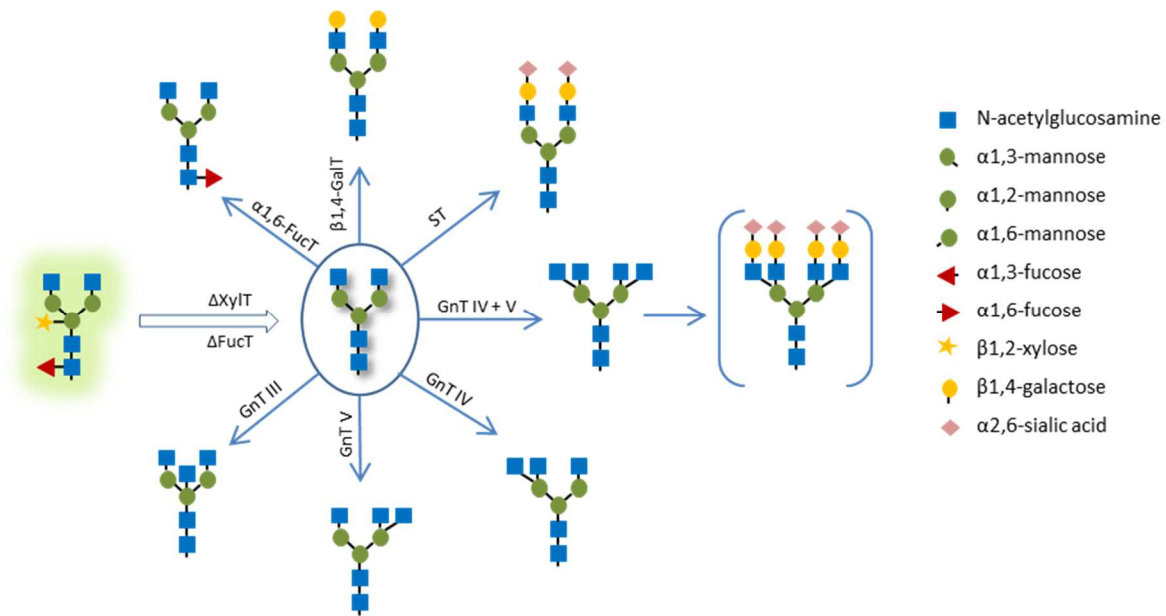


Figure 3. Engineering the plant N-glycosylation pathway. Overview of all the effort dedicated to humanize plant N-glycans by knockout of endogenous transferases (hollow arrows, the inactivated enzyme is preceded by the symbol 'Δ') and knock in of heterologous transferases (full arrows). A typical immunogenic plant complex N-glycan structure is highlighted in green on the left. Inactivation of the genes coding for β1,2-xylosyltransferase and core α1,3-fucosyltransferase yields a non-immunogenic structure with terminal GlcNAcs, indicated by a blue circle, that forms the basic structure for almost all further manipulations. Although not depicted here, also the Lewis X epitope has been introduced *in planta*. The multiantennary, sialylated glycan structure is placed between brackets as it has not been produced as such yet. Abbreviations used: XylT, β1,2-xylosyltransferase; FucT, core α1,3-fucosyltransferase; α1,6-FucT, α1,6-fucosyltransferase; β1,4-GalT, β1,4-galactosyltransferase; GnT I to V, N-acetylglucosaminyltransferase I to V; ST, set of enzymes responsible for the biosynthesis, activation, transport and transfer of sialic acid onto N-glycans. All glycan structures are depicted according to the recommendations of the international consortium of functional glycomics⁷.

⁷ www.functionalglycomics.org

A key for each lock, towards the desired glycoform

Ever since the conception of PMPs and the discovery of immunogenic plant N-glycans, major efforts have been dedicated towards ‘humanizing’ the plant N-glycosylation pathway. At the very least this meant silencing or knocking out the genes coding for fucosyl and xylosyltransferase combined with the introduction of somewhere from one up to no less than six recombinant enzymes to obtain more human-like N-glycans devoid of xylose and fucose (see previous section). All the work done to engineer the plant N-glycosylation pathway is justified by the fact that different glycans are required to bestow different characteristics onto a therapeutic protein. For example, glycoforms of recombinant erythropoietin for human use with sialylated, multi-antennary N-glycans would be highly desirable because these large structures would overall significantly increase the size of the otherwise small erythropoietin (Nagels et al., 2012a). This would decrease the clearance of erythropoietin by the kidneys and result in an increased circulation half-life and hence less frequent administration would be needed.

Another good example are the immunoglobulins that carry two bi-antennary N-glycans on their Fc domain (Figure 2B, glycan 5 and 6). These N-glycans do not interfere with antigen recognition but determine the three-dimensional shape of the Fc domain and by doing so, modulate binding to activating and inhibitory Fc receptors, thus influencing the biological activity of the IgG (Buck et al., 2013; Shields et al., 2001; Umana et al., 1999). Most glycoforms of human IgG carry N-glycans with terminal galactoses and it has been shown that removal of core fucose attached to the innermost GlcNAc enhances antibody dependent cellular cytotoxicity (ADCC), an important effector function of IgGs (Shields et al., 2002). Interestingly, this observation was made for both α 1,3 as α 1,6-fucose, respectively from plant and mammalian origin (Forthal et al., 2010). Moreover, sialylated IgGs, which represent less than 10 % of all glycoforms, were shown to have anti-inflammatory properties (Anthony et al., 2008; Kaneko et al., 2006; Nimmerjahn & Ravetch, 2007). The N-glycans on the Fc-domain are thus a crucial determinant of the biological activity of any given antibody.

All the above mentioned examples concern cases where the presence of specific glycan groups is advantageous. The opposite also holds true and in some cases the absence of glycans, such as immunogenic plant glycans, is beneficial. This idea can be further expanded to situations where it would be interesting to avoid a glycan structure entirely. Exposed viral proteins can be heavily glycosylated and it is believed that this serves to shield sensitive epitopes from the immune system, a phenomenon known as ‘glycan shielding’ (Vu et al., 2011). Vaccines based on viral proteins lacking these glycans might be more efficacious (Wang et al., 2009). In another interesting line, for some monoclonal antibodies that bind and eliminate a factor circulating in the blood, effector functions are

undesirable and could provoke unwanted side effects, such as is the case for e.g. the therapeutic antibody infliximab that binds the potent proinflammatory cytokine tumor-necrosis-factor alpha (TNF- α). Removing the N-glycans on the Fc domain changes the 3D conformation impeding interaction with activating and inhibitory Fc receptors and hence abolishing effector functions (Buck et al., 2013; Shields et al., 2001).

Taken together, there is a demand for deglycosylated variants of specific therapeutic glycoproteins. This could simply be achieved by amino acid substitution of the asparagines to which the N-glycans are attached. Nevertheless, this would change the primary structure of the recombinant protein, possibly negatively influencing its desired therapeutic characteristics. Moreover, N-glycans play a pivotal role in the correct folding of a glycoprotein in the ER by attracting chaperones and oxidoreductases and interacting with the quality control mechanism (Pearse & Hebert, 2010). This generates a paradox, where N-glycans on a therapeutic are both necessary and unwanted.

A solution to this problem would be to produce a therapeutic glycoprotein and then treat it enzymatically with an endoglycosidase. However, this adds a number of steps to the entire production process, which would further increase its complexity. Recently this process was simplified by combining these steps *in vivo* via the transient co-expression of glycoproteins together with the bacterial endoglycosidase PNGaseF in *Nicotiana benthamiana* (Mamedov et al., 2012; Mamedov & Yusibov, 2013). The goal of these experiments was to produce a non-glycosylated version of a vaccine candidate from the malaria parasite *Plasmodium falciparum*. The antigen carries several putative N-glycosylation sites to which glycans are attached upon its recombinant expression in plants. However, the native protein is not glycosylated and the presence of the N-glycans interferes with its correct folding. Co-expression together with the PNGaseF effectively deglycosylated the antigen that was significantly better detected by monoclonal antibodies raised against the native antigen than its glycosylated counterpart (Mamedov et al., 2012).

Although this system is undoubtedly useful on specific occasions, for most proteins (those that are glycosylated in their native form) the innermost GlcNAc, which is cleaved off by PNGaseF, is an integral structural element, because of its direct contact with the peptide backbone, and is required to maintain the protein's native fold. Release of this GlcNAc by PNGaseF will alter the conformation of the protein and this can result in loss of stability and degradation of the protein. For example, full truncation of the glycans present on the Fc-domain of IgGs has been shown to increase susceptibility to the protease papain, drastically lower the thermal stability and increase the propensity to aggregate (Buck et al., 2013; Latypov et al., 2012; Mimura et al., 2000).

Engineering the *Arabidopsis* seed platform, Glycodelete Technology™

Rationale

Up to date, no production platform exists that can efficiently cope with the removal of the glycans on a native glycoprotein without affecting its overall fold. Therefore, we set out to glycoengineer the *Arabidopsis* seed-platform to expand its capacities towards the production of glycoproteins with a homogeneous glycoprofile that could show reduced immunogenicity and possibly enhanced therapeutic efficacy when minimally glycosylated.

In an interesting line of work, mammalian cells were engineered to produce glycoproteins with short, truncated glycans only. This ingenious strategy referred to as ‘GlycoDelete Technology’ was conceived with the intention of reducing the heterogeneity that is associated with complex N-glycans and that is detrimental for the downstream processing of a therapeutic protein, but without interfering with the folding-catalyzing function of N-glycans in the ER (Meuris et al., manuscript in press by Nature Biotechnology). The authors hypothesized that a single endoglycosidase delivered to the Golgi would be able to hydrolyse the glycans present on therapeutic proteins between the two GlcNAc residues of the chitobiose core. The single GlcNAc residue still attached to the protein backbone would then serve as substrate for the endogenous galactosyltransferase and sialyltransferase thus generating short trisaccharides with terminal sialic acids as is encountered on the larger complex glycans. The trisaccharides mimic the the latter, but without the associated heterogeneity and this is a desirable feature for most therapeutic proteins. Indeed, when a fungal endoglycosidase was targeted to the *trans*-Golgi of a cell line producing high-mannose N-glycans, it completely trimmed the glycans present on a set of therapeutic reporterproteins that were coexpressed. These truncated glycans were then effectively further elongated with galactose and sialic acid, albeit to various degrees. Despite the fact that some heterogeneity remained due to the concurrent presence of mono-, di- and trisaccharides, the overall glycan heterogeneity had been considerably reduced. Moreover, a neutralizing monoclonal antibody (anti-CD20) produced in the ‘GlycoDelete’ cells showed enhanced serum half-life and better safety profile because the truncated glycans on its Fc-domain interfered with binding to activating an inhibitory Fc-receptors, hence abolishing unwanted effector functions.

Based on these previous results validated in mammalian cells we implemented the GlycoDelete concept in *A. thaliana* to tackle the issue of complex N-glycan heterogeneity and immunogenicity. To achieve this goal, an *Arabidopsis* line has to be created that allows the addition of N-glycans to a glycoprotein in the ER, so that it can reach its correct fold, and delivers the protein in a minimally

glycosylated state after passage through the secretory pathway. In accordance with the mammalian GlycoDelete technology, a recombinant endoglycosidase present in the Golgi apparatus should be able to digest a glycoprotein passing through the secretory pathway (Figure 4). The glycoprotein would only be released from the ER upon reaching its correct fold and hence this system would deliver a correctly folded, minimally glycosylated protein. The different components of this system are discussed in the following paragraphs.

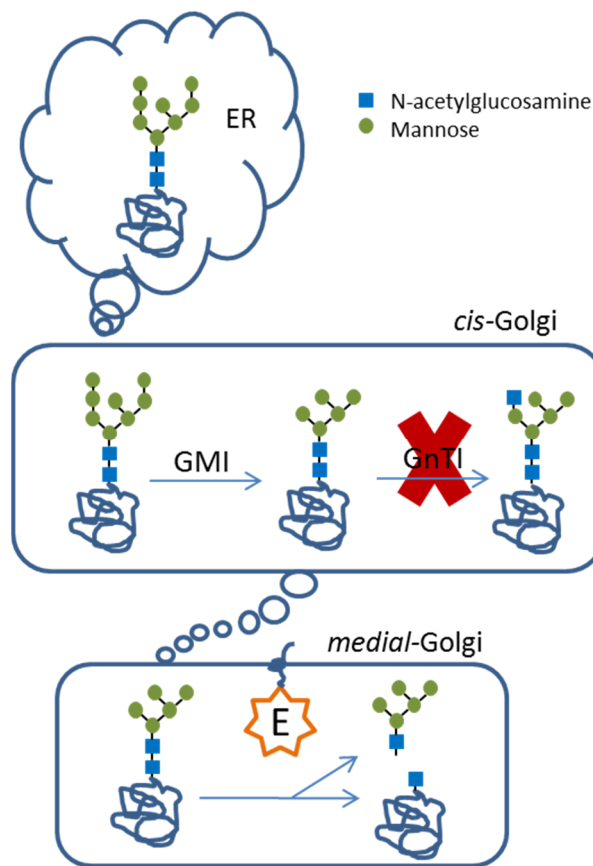


Figure 4. The GlycoDelete concept implemented *in planta*. An endoglycosidase, more specifically the endoT from *Trichoderma reesei* (Stals et al., 2010), is expressed in a background lacking GnTI activity (i.e. the *Arabidopsis thaliana* complex-glycan-deficient mutants that are GnTI knockout lines (von Schaewen et al., 2008; von Schaewen et al., 1993)), such that proteins passing through the secretory pathway will be decorated with high-mannose glycans only, which are the substrate of the endoT. This results in the trimming of these glycans to completely homogenous, minimal structures, consisting of a single GlcNAc that lack possibly immunogenic plant glycans while maintaining the native protein fold. Abbreviations used: ER, endoplasmic reticulum; GMI, Golgi-mannosidase-I; GnTI, N-acetylglucosaminyltransferase-I; E, endoT.

***The Trichoderma reesei* Endo-N-acetyl- β -D-glucosaminidase**

The *Trichoderma reesei* endo-N-acetyl- β -D-glucosaminidase (endoT, Figure 5) was selected as tool to glycoengineer the *Arabidopsis* seeds. It was discovered as an extracellular component in the growth medium of the fungus *Trichoderma reesei*, also known as *Hypocrea jecorina* (Stals et al., 2010); the different names refer to the asexual and sexual phase of the fungus respectively. It is an important industrial microorganism renowned for its extraordinary capacity of secreting large quantities of biomass-degrading enzymes, which can amount up to 100 grams per liter for some industrial strains

(Druzhinina et al., 2011; Martinez et al., 2008). Enzymes excreted by *Trichoderma* carry truncated N-glycans consisting of a single GlcNAc attached to the amino acid backbone and it was discovered that the endoT is responsible for this activity (Stals et al., 2010).

Purification of the enzyme from the growth medium and subsequent characterization revealed several interesting features; the active form is both N and C-terminally processed releasing peptides of respectively 9 and 46 amino acids long, endoT has a substrate specificity for high-mannose, but not complex N-glycans and cleaves the bond between the innermost and second GlcNAc moieties of the glycan leaving one GlcNAc residue attached to the protein backbone. Three N-glycan attachment sites are present of which two are occupied with a single GlcNAc suggesting auto-cleavage. The enzyme is both structurally and functionally related to other endoglycosidases (Stals et al., 2012; Stals et al., 2010).

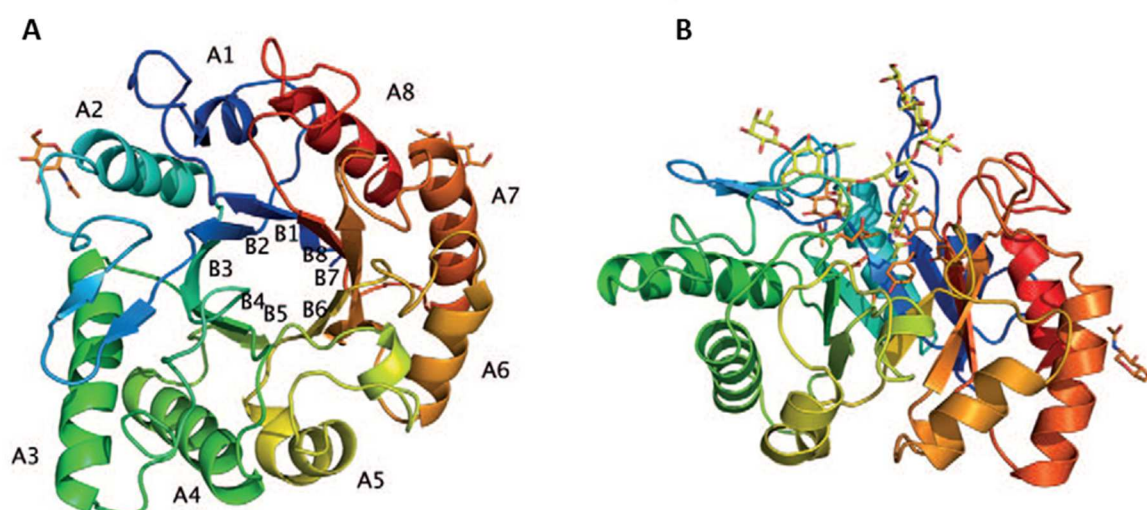


Figure 5. Structure of the *Trichoderma reesei* endoT. Figure taken from Stals et al. (2012). Crystallized endoT was shown to have a complete $(\beta/\alpha)_8$ -TIM barrel fold. The structure is rainbow colored according to the residue number, starting with blue at the N-terminus and ending with red at the C-terminus. The single GlcNAc residues at position N70 and N240 resulting from auto-deglycosylation are represented as orange-colored stick models. (A) Upper view of the endoT with indication of the α -helices and β -sheets that make up the barrel fold. (B) Side view of the endoT with an octasaccharide modeled into the active site to indicate its position.

An advantage of endoT in the glycoengineering context is that, in contrast to other endoglycosidases that completely remove the N-glycan, endoT releases the N-glycan while leaving the proximal GlcNAc

residue attached to the protein. This is of considerable significance since this residue is in direct contact with the protein backbone and forms an integral part of the overall structure of the protein. EndoT activity thus eliminates the glycan structure while leaving the structurally important proximal GlcNAc in its place, hence conserving the native protein fold.

The Arabidopsis thaliana complex-glycan-deficient mutant

To ensure optimal deglycosylation, the *Arabidopsis thaliana complex-glycan-deficient (cgl)* mutant was selected as host for the glycoengineering. The *cgl* mutant was discovered during a chemical mutagenesis screen and solely carries N-glycans of the high-mannose type, with a predominance of Man₅, which are an excellent substrate for the endoT (Stals et al., 2010; von Schaewen et al., 1993). A deficiency in the N-acetylglucosaminyltransferase I (GnTI) that initiates the conversion of high-mannose to complex N-glycans, lies at the base of this observed glycan profile (Gomez & Chrispeels, 1994; von Schaewen et al., 1993). GnTI catalyzes the transfer of a terminal GlcNAc to the α 1,3 branch of Man₅ in the *cis*-Golgi and its product is a prerequisite for all further maturation steps (Wenderoth & von Schaewen, 2000).

Analysis showed that point mutations are responsible for the lack of GnTI function in both allelic *cgl* variants (Frank et al., 2008). In the *cgl1-1* mutant (Nottingham Arabidopsis stock centre number N6192) this point mutation leads to the introduction of an additional N-glycan in the active site of the enzyme, impeding it of reaching its correct fold (Strasser et al., 2005). The resulting enzyme is inactive and retained in the ER, however, conditions that stimulate underglycosylation can partially lift the block in conversion of high-mannose to complex N-glycans because in absence of the N-glycan attached to it, the asparagine introduced by point mutation in the enzyme's active site still allows for residual GnTI activity (Frank et al., 2008; Strasser et al., 2005). The *cgl1-1* mutant thus represents a conditional mutant and it has recently been reported that recombinant glycoproteins expressed in its seed carried predominantly high-mannose N-glycans, but with a significant fraction of complex N-glycans, arguing for some GnTI activity during seed development (He et al., 2012a; He et al., 2012b; He et al., 2013).

In contrast to the *cgl1-1* mutant, the *cgl1-2* is a true null mutant (Nottingham Arabidopsis stock centre number N16366). Here, a point mutation at the end of an intron splice site creates a frame shift that introduces a premature stop codon and no residual enzymatic activity can be found (Frank et al., 2008). Both *cgl* mutants perform similar to wild type *Arabidopsis* (Col-0) under normal conditions and specific heat, cold and light stress although they are possibly more susceptible to

stray pathogen infection and, except for longer root hairs, they display no obvious phenotype (Frank et al., 2008; von Schaewen et al., 1993).

Although the glycans present on proteins produced in the *cgl* mutants are of the high-mannose type and are non-immunogenic by nature, they are still considerably heterogeneous, ranging from Man₅ to Man₈, and give rise to different glycoforms, which is an undesirable feature for any therapeutic protein. Additionally, high-mannose glycans can target a therapeutic protein to a specific subset of cells that carry the mannose receptor, which results in the rapid clearance of the therapeutic from the blood stream. All in all, the presence of these high-mannose glycans on a therapeutic protein should, except for specific applications, be avoided.

Chapter IV

Glycoengineering in *Arabidopsis* seeds, results & discussion.

Robin Piron, Francis Santens, Annelies De Paepe, Julie Goossens, Els Van Lerberge, Jonah Nolf, Ann Depicker and Nico Callewaert.

R.P. wrote the chapter. R.P., F.S., A.D. and N.C. conceived the experiments. R.P., F.S., A.D.P., J.G., J.N. and E.V.L. performed the experiments. A.D.P. cloned the ASP1 gene constructs and introduced it into wild type seeds. J.G. and E.V.L. helped with the screening and propagation of the transgenic lines. J.N. assisted during ASP1 purification and F.S. during glycan analysis. A.D. and N.C. edited the chapter.

A reporter protein for the GlycoDelete technology, the activation-associated secretory protein 1

The activation-associated secretory protein 1 (ASP1, Figure 1) from the helminth *Ostertagia ostertagi* was selected to serve as reporter glycoprotein to probe the functionality of the GlycoDelete technology implemented *in planta*. The organism is the major parasitic nematode of cattle in the temperate regions of the world and poses a serious problem (Geldhof et al., 2002). Treatment is based on anti-helminthic drugs but because of concerns about resistance to these drugs and their residual presence in consumer goods, e. g. milk and meat, alternative measures for control, such as vaccination, are being investigated. So far, excretory/secretory proteins from adult larvae that were purified by thiol-sepharose chromatography were shown to provide the strongest protective immunity (Geldhof et al., 2002; Geldhof et al., 2000; Geldhof et al., 2004). Further analysis revealed that ASP1 and ASP2 are the predominant components of this ES-thiol fraction and purified ASPs elicited the same response as the crude fraction, identifying them as the main immunogens (Geldhof et al., 2003; Meyvis et al., 2007).

However, isolation of large quantities for vaccination is difficult and there is great interest to produce recombinant versions of these antigens. ASPs have been produced in the yeast *Pichia pastoris* and in baculovirus-infected insect cells, but while very high antibody titers were elicited, neither recombinant version was able to provide any protection whatsoever *in vivo* and the antibodies raised in calves against these ASPs were not able to recognize the native ASP, pointing towards structural differences between the recombinant and native ASPs (Geldhof et al., 2008; Van Coppennolle et al., 2012). The crystal structure of Pichia-produced ASP1 has recently been elucidated and reveals that the antigen consists of an N-terminal domain with an α - β - α fold, a typical feature for ASPs, and a C-terminal hinge region that allows for dimerization, as occurs *in vivo* (Borloo et al., 2013; Meyvis et al., 2007). Six internal disulfide bridges are present that stabilize the monomer and a peculiar feature is that one of these links the C to the N-terminus, thus generating a cyclic structure. The seventh bridge connects both monomers and it is postulated that its location in the less rigid hinge region allows the dimer to be quite flexible. This is probably related to the *in vivo* function of the ASPs in the host-parasite interaction although their exact role still awaits elucidation (Borloo et al., 2013).

The two N-glycan attachment sites of both ASPs are occupied *in vivo* with predominantly (90 %) paucimannosidic and hybrid glycans that are core α 1,6-fucosylated, but their function is still unclear (Meyvis et al., 2008). Antibodies raised against the native ASP recognize the ASP in a glycan-independent manner. However, reduction and denaturation of the ASP drastically reduced antibody binding, revealing the importance of conformational epitopes of the protein backbone

(Meyvis et al., 2008). As native ASP is purified through a thiol-sepharose chromatography step, the purified product is a disulfide-scrambled mixture, in contrast to the recombinant proteins produced in *P. pastoris* and insect cells. It is possible that this disulfide scrambling exposes epitopes to which antibodies are raised that are important in the vaccine-induced protection. Anyhow, the recombinant ASP1 from *Pichia* and baculovirus-infected insect cells apparently lack some structural element, rendering them unprotective, and this spurred the production of recombinant ASP1 in the seeds of *Arabidopsis thaliana* (unpublished results, Dr. A. De Paepe). It was hypothesized that the seed platform might just be capable of correctly incorporating this enigmatic structural element into the seed-produced ASP1 and bestow it with protective properties. The ASP1 could indeed be produced in the seeds of *Arabidopsis*, however, competitive ELISA showed that the seed-produced ASP1 is not recognized by neutralizing ASP1 antibodies, what seems to indicate that the seed-produced ASP1 also lacks the structural element that is absent in the other recombinant ASP1s.

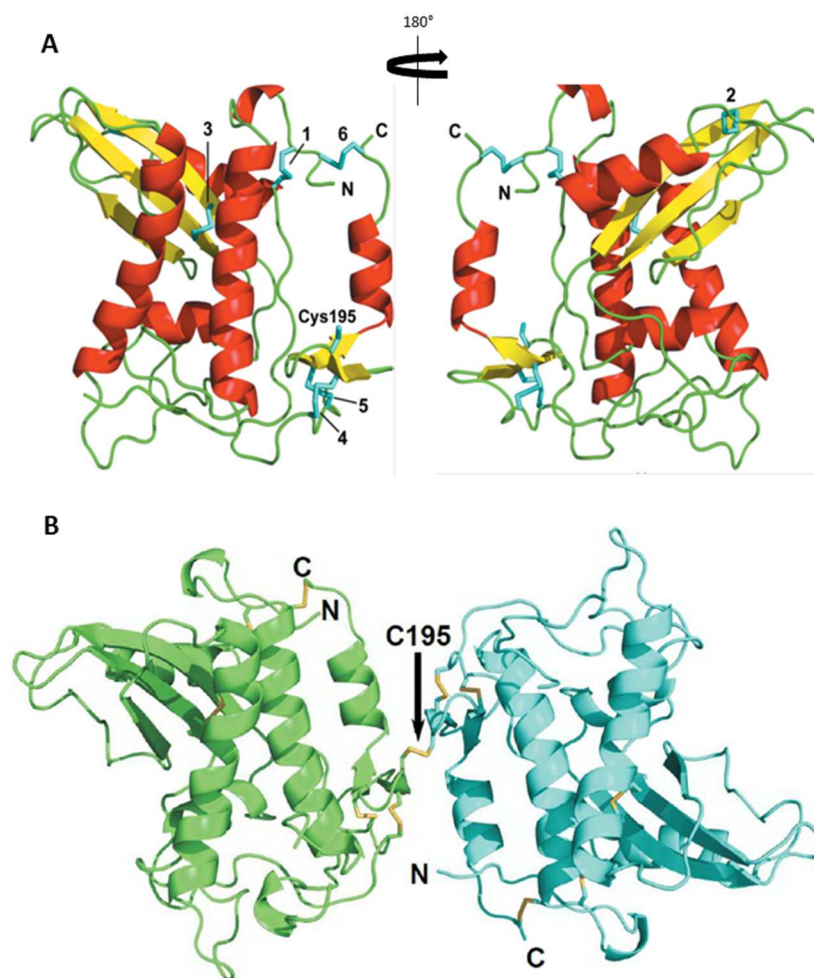


Figure 1. Structure of the *Osteragia ostertagi* ASP1. Figure adapted from Borloo et al. (2013). Recombinant ASP1 was produced in *P. pastoris*, crystallized and its structure elucidated. (A) Cyclic ASP1 monomer with its secondary structure highlighted and six internal cysteine bridges indicated. Cys 195 is involved in dimerization of the ASP1 monomers. Alpha-helices and β -sheets are depicted in red and yellow respectively. (B) ASP1 dimer connected by the disulfide bridge formed by both Cys 195 residues.

Results

EndoT expression in Arabidopsis thaliana complex-glycan-deficient seeds

A first step towards the generation of an *Arabidopsis*-seed platform capable of producing correctly folded therapeutic proteins with truncated N-glycans is the coordinated integration of the endoT into the secretory pathway of the seeds of the *Arabidopsis cgl* mutant. To avoid interference with the folding of nascent proteins in the ER, the endoT needs to be targeted to the Golgi apparatus and this in a seed-specific way. To this goal, the signal sequence of the endoT was swapped for that of the *Arabidopsis* β 1,2-xylosyltransferase which resides in the stacks of the *medial*-Golgi. The xylosyltransferase is a type-II integral membrane protein, i.e. with an N-terminal cytoplasmic tail that is 10 amino acids long, a single-span transmembrane domain and C-terminal catalytic domain facing the lumen of the ER (Strasser et al., 2000). The first 32 amino acids that make up the cytoplasmic tail and transmembrane domain were shown to be sufficient for correct localization in the *medial*-Golgi (Pagny et al., 2003) and these were used to substitute the endoT signal sequence (Figure 2B). The construct was cloned into a seed-specific expression cassette under control of the *β -phaseolin* promoter the *arcelin5-1* terminator and used to transform the *Arabidopsis cgl1-1* and *1-2* mutants (Figure 2A). T1 transformants were selfed to obtain T2 segregating seed stocks which were analyzed by western blot for the presence of endoT (Figure 3).

A similar pattern of endoT expression can be observed in both allelic mutant backgrounds (Figure 3A and 3B), typified by a doublet band around 40 kDa and another major band around 30-33 kDa that is surrounded by three additional bands of lower intensity of which one is situated higher at about 37 kDa and the remaining two lower in the 15 – 25 kDa range. The highest band of the doublet corresponds to a glycoform of the lower band as it disappears after digestion with PNGaseF (see later) and the latter probably represents the unprocessed endoT, as its observed size corresponds with that of the full length, unglycosylated endoT (41.3 kDa, Isx:endoT; Table 1). The other major band around 30 – 33 kDa seems to coincide with the size of the endoT catalytic domain (31.7 kDa) as it is encountered in the extracellular medium of its host, the fungus *Trichoderma reesei*. *In vivo*, the active extracellular endoT is the product of proteolytic processing at its N and C-terminus, releasing peptides of 1.1 and 4.7 kDa respectively. These propeptides are present in the construct that was used to transform the *Arabidopsis cgl* mutants and it is thus possible that the endoT is processed in a similar way in *Arabidopsis cgl* seeds as occurs *in vivo*. In accordance with this hypothesis, the lower doublet band could represent non glycosylated and unprocessed endoT (Isx:endoT, 41.3 kDa), the highest low intensity band around 37 kDa may represent the C-terminally processed endoT (36.6 kDa) and the major band 30 – 33 kDa would then be the fully processed endoT catalytic domain (31.7

kDa). The two remaining low intensity bands could be the result of degradation. An important remark to be made is that if full processing indeed occurs in the seed, then the *lsx* will be released from the *endoT* catalytic domain as an integral part of the N-terminal propeptide and this would convert the *endoT* into a secretory protein instead of retaining it in the *medial*-Golgi.

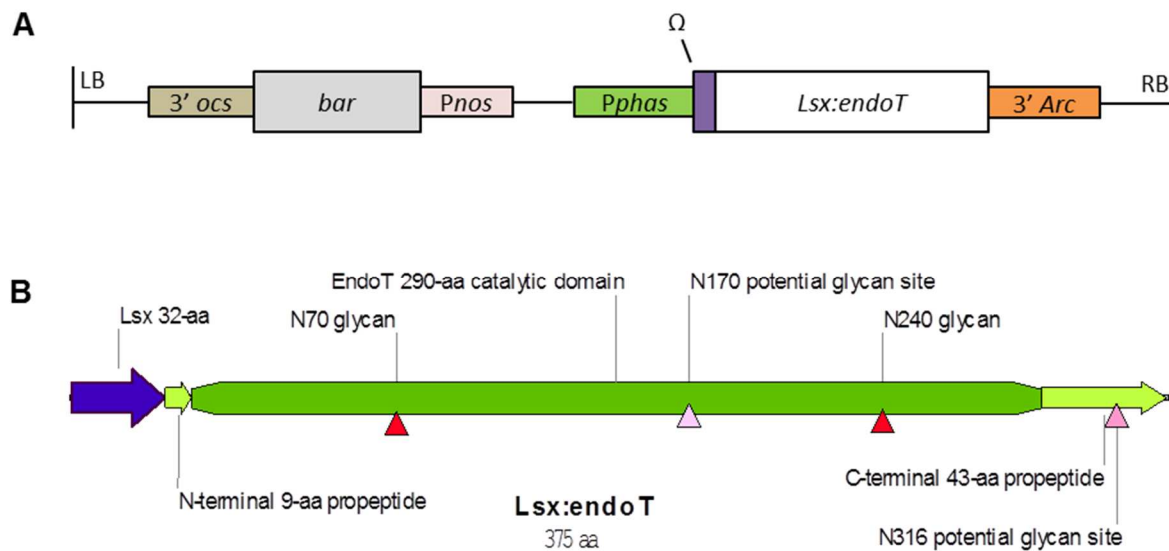


Figure 2. T-DNA bearing the *lsx:endoT* gene and translation thereof. (A) The T-DNA from the binary vector used to transform *Arabidopsis*, carrying the seed-specific expression cassette in which the *endoT* construct was cloned and the *bar* gene to convey phosphinothricin (ppt) resistance to transformed seedlings. The *lsx:endoT* gene is placed under control of the *β-phaseolin* promoter and the *arcelin5-I* terminator, both derived from seed storage proteins of *Phaseolus vulgaris* (common bean), to promote high, seed-specific expression. (B) The *lsx:endoT* fusion protein, consisting of the unprocessed *endoT* without ER signal sequence coupled at its N-terminus to the *medial*-Golgi targeting signal of the *Arabidopsis* β1,2-xylosyltransferase (*lsx*). Given are the fully processed *endoT* catalytic domain, both N and C-terminal propeptides, the *medial*-Golgi targeting signal, the two sites that carry single GlcNAc residues *in vivo* (red triangle) and the two remaining potential N-glycan attachment sites (pink triangles). Abbreviations used: aa, amino acids; *Pnos*, nopaline synthase gene promoter; 3'ocs, octopine synthase gene terminator; *bar*, bar gene conveying phosphinothricin resistance; *Pphas*, β-phaseoline gene promoter; Ω, leader sequence of the 'tobacco mosaic virus'; 3'Arc, 600 basepairs arceline5-I gene terminator; *lsx:endoT*, gene construct coding for the *lsx:endoT* fusion protein; *lsx:endoT*, translational fusion between the β1,2-xylosyltransferase *medial*-Golgi targeting signal and the N-terminus of the *endoT* without signal sequence.

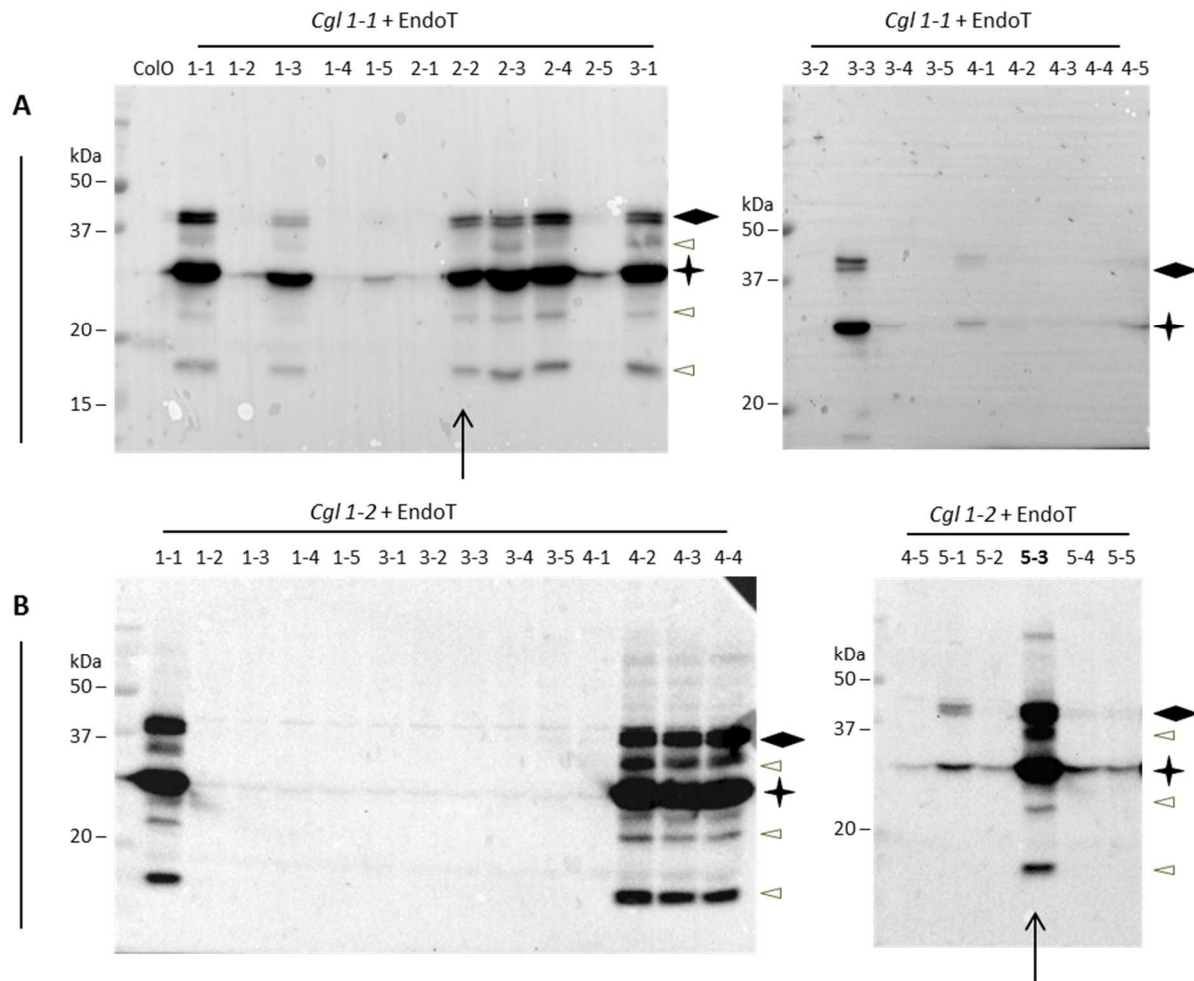


Figure 3. Expression analysis of *endoT* in the *Arabidopsis cgl 1-1* and *1-2* mutant. Extracts of T2 segregating seed stocks of *Arabidopsis cgl 1-1* (A) and *cgl 1-2* (B) transformed with the *endoT* construct were analyzed by western blotting using a polyclonal rabbit α -*endoT* serum (1:5000). Gels were run under reducing conditions and 20 μ g of protein was loaded. T-DNA single locus integration lines that were selected to be propagated to homozygosity are marked with an arrow. A diamond indicates the doublet band, whereas the major band around 30 kDa is represented by a star and the three additional bands of minor intensity by hollow arrowheads. Abbreviations used: kDa, kilodalton; *cgl*, *Arabidopsis complex-glycan-deficient* mutant; ColO, *Arabidopsis thaliana* ecotype Columbia.

Table 1. Physical parameters of the recombinant endoT

	Amino acids	Glycosylation sites	Predicted Mw (kDa)	
			min	max
Lsx:endoT	374	3	41,3	47,3
EndoT propeptide	342	4	37,4	45,4
EndoT catalytic domain	290	3	31,7	37,7
lsx	32	0	3,9	3,9
N-terminal fragment	9	0	1,1	1,1
C-terminal fragment	43	1	4,7	6,7

Table 1. Characteristics of the recombinant endoT. For any peptide, the number of amino acids, number of N-glycosylation sites and the minimum and maximum size are given (minimum size plus 2 kDa multiplied by the number of potential glycosylation sites). EndoT propeptide and catalytic domain correspond to the un- and fully processed endoT. Abbreviations used: Mw, molecular weight; kDa, kilodalton; pI, isoelectric point; Lsx, Arabidopsis β 1,2-xylosyltransferase *medial*-Golgi targeting signal; Lsx:endoT, endoT fused at its N-terminus to the Lsx.

After segregation analysis the *cg1-1* endoT 2-2 and *cg1-2* endoT 5-3 were identified as the single locus T-DNA integration lines with the highest endoT expression and these were subsequently propagated to homozygosity. To further elucidate the endoT expression pattern, seed extracts of both homozygous endoT lines were subjected to a PNGaseF treatment and analyzed via western blotting (Figure 4). This revealed that only the most upper band is susceptible to glycan digest and results in the collapse of the doublet band showing that it consists of two glycoforms of the same protein (black diamond, Figure 4). Again, no difference could be observed between both allelic mutant backgrounds. An interesting detail that is also revealed by this blot is that the intensity of the endoT pattern is different for the extracts with and without the detergent dodecylmaltoside. The latter is used to solubilize the endoT from the membrane in which it is anchored through its N-terminal Golgi localization signal. Hence, the doublet band and the low intensity band right below it that are only faintly stained in the extracts without detergent probably represent membrane-anchored variants of the endoT. In accordance with the hypothesis postulated in the previous section, the doublet band then represents two glycoforms of the unprocessed (Lsx:endoT), membrane-anchored endoT. The fainter band right below the doublet then corresponds to partially processed endoT and its lower intensity in the detergent-free extract would suggest that it is membrane-anchored and thus C-terminally, but not N-terminally processed. Finally, the equal intensity of the band at around 33 kDa in both types of extract points towards loss of the membrane

anchor and hence, this represents the fully processed endoT catalytic domain only, which is confirmed by comparison to the catalytic domain of yeast-produced endoT (see later).

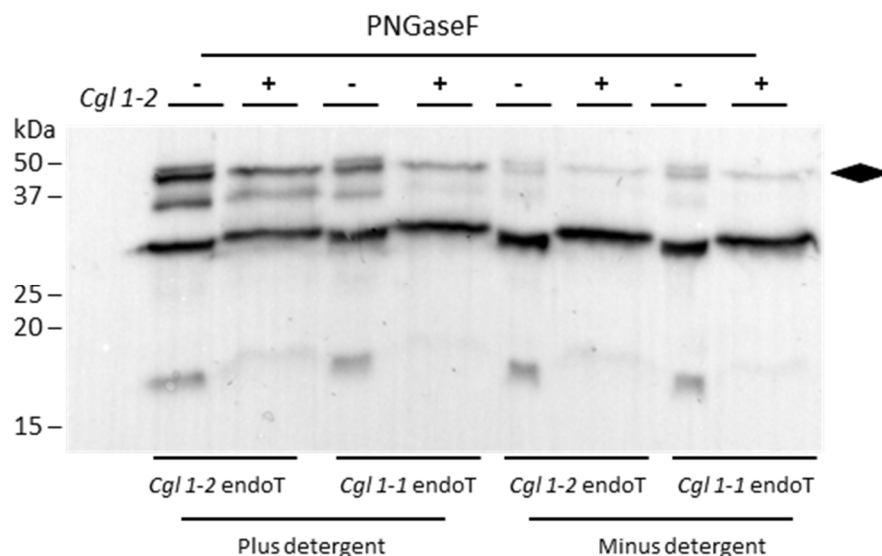


Figure 4. Glycan analysis of the endoT produced in *cgl* seeds. Seed stocks of *cgl* lines T3 homozygous for endoT (*cgl 1-1* endoT 2-2-7 and *cgl 1-2* endoT 5-3-1) were extracted at a fixed w/v ratio, with or without the detergent dodecylmaltoside. The resulting extracts were digested with PNGaseF and analyzed via western blot using a polyclonal rabbit α -endoT serum (1:5000). 10 μ l of seed extract was incubated with PNGaseF, loaded onto the gel and run under reducing conditions. The collapsed doublet band is indicated with a black diamond. Abbreviations used: kDa, kilodalton; *cgl*, *Arabidopsis* complex-glycan-deficient mutant.

***EndoT* deglycosylates endogenous seed glycoproteins**

After transformation of the endoT in the *cgl* mutants and confirmation of its presence in the seeds, the next step towards a functional GlycoDelete platform is to show enzymatic activity. EndoT has a substrate specificity for high-mannose N-glycans, which are the predominant, if not only glycans present on proteins in both allelic *cgl* backgrounds. It was hypothesized that if the seed-produced endoT would show enzymatic activity, then it should also digest the endogenous seed glycoproteins. To study the glycans present on these proteins, crude seed extracts were prepared from *Arabidopsis* ecotype Columbia, both *cgl* mutants and low to high endoT expressing lines. These were separated by SDS-PAGE and visualized by western blotting using the lectin Concanavalin-A (ConA) coupled to horse radish peroxidase (HRP). ConA binds high-mannose N-glycans and together with the HRP enzymatic moiety, this allows for chemiluminescent detection of high-mannose bearing glycoproteins. A first western blot shows the endoT content of the selected lines ranging from

non-detectable over intermediate to high (from ‘-’ over ‘+’ to ‘++’ respectively, Figure 5A). The second western blot reveals the high-mannose bearing proteins present in the extracts (Figure 5B). In the Columbia extract, where N-glycans can mature into hybrid and complex glycans, only a few bands are visible (Figure 5B). As expected, in the *cgl* mutants where high-mannose glycans are the sole glycoforms present, much more proteins are stained. However, when comparing the profile of both mutants to transformed lines expressing endoT, it immediately becomes clear that endoT expression reduces this pattern and that there is an inverse relation between the endoT content of the extracts and the number of bands detected with only two major bands remaining in the high and homozygous endoT lines (‘++’ and ‘H’ respectively, Figure 5A & 5B). To confirm that the patterns revealed by ConA detection (Figure 5B) truly represent high-mannose bearing glycoproteins, the western blot was performed in duplicate and an excess of soluble ConA ligand (methyl- α -D-mannopyranoside) was added during the incubation of the blot with ConA to specifically inhibit the glycan-binding site of ConA (Figure 5C) showing that binding of the lectin was indeed through its mannoside binding site, except for a remaining band around 20 kDa.

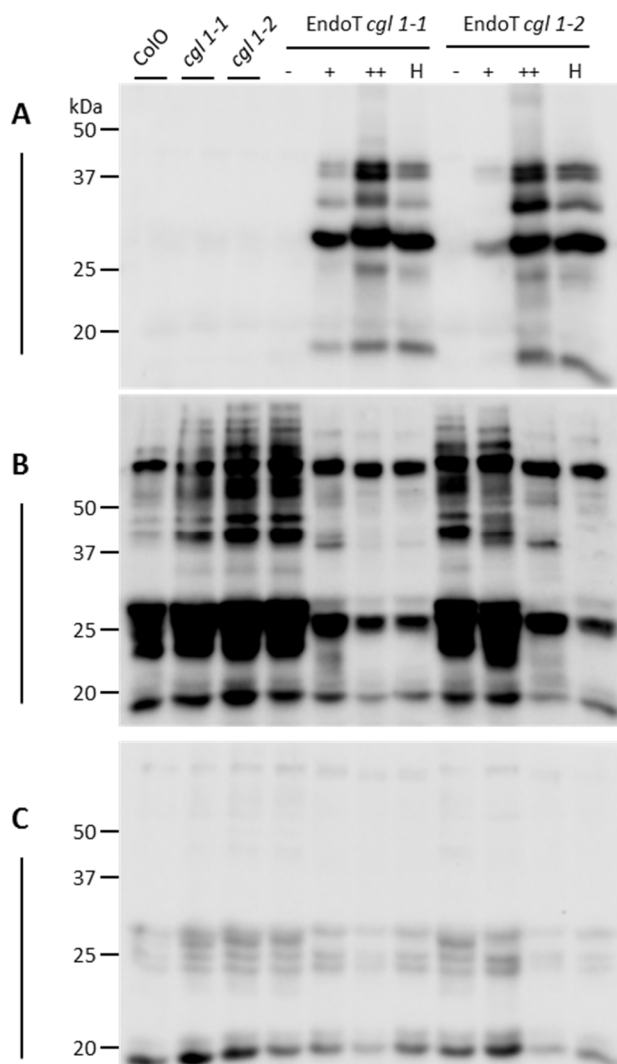


Figure 5. EndoT deglycosylates endogenous seed proteins. Seed extracts from *Arabidopsis* ecotype Columbia, both *cgl* mutants and endoT expressing lines were prepared at a fixed w/v ratio and analyzed by western blot using (A) a polyclonal rabbit α -endoT serum to assay the endoT content of the different lines, (B) ConA to detect high-mannose bearing glycoproteins and (C) ConA in the presence of excess soluble ligand to reveal binding of ConA through other sites than its mannoside binding site. 10 μ l of extract was loaded and the gels were run under reducing conditions. Abbreviations used: Col-0, *Arabidopsis* ecotype Columbia; *cgl*, *Arabidopsis* complex-glycan-deficient mutant; -/+ / ++, non-detectable, intermediate and high endoT content respectively; H, homozygous.

ASP1 expression in wild type, *cgl* and glycodelete™ *Arabidopsis* seed

The reduction in ConA-sensitive seed glycoproteins in *cgl* lines expressing endoT delivered the proof-of-concept of the activity of the recombinant endoT. This left one final question to be addressed to demonstrate the full utility of the GlycoDelete seed-platform, namely that a foreign glycoprotein would also be processed by the endoT upon coexpression in the seed. To tackle this issue, the ASP1 from the parasitic nematode *Ostertagia ostertagi* was selected to be introduced in the endoT expressing *cgl* lines. For both the *cgl 1-1* as the *cgl 1-2* mutant a line homozygous for endoT was available (EndoT 2-2-7 and 5-3-1 respectively, propagated from line 2-2 and 5-3, Figure 3) and although both mutants share the same phenotypical characteristics, they are genotypically different and represent a conditional and a true null mutant respectively. To ensure complete absence of GnTI functionality and complex plant N-glycans, we opted to continue only with the *Arabidopsis cgl 1-2* mutant homozygous for endoT (EndoT 5-3-1, propagated from line 5-3, Figure 3) in further experiments. From here on we will refer to the latter as the GlycoDelete seed-platform.

A secretory version of ASP1 was N-terminally fused to the ER signal sequence of the *Arabidopsis* 2S2 seed storage protein and cloned into a seed-specific expression cassette under control of the same regulatory elements that were used to drive expression of the endoT, except that the short version of the *arceline5-1* 3'-UTR, 600 basepairs (bps), was substituted for the full length 3'-UTR (4000 bps; unpublished results, Dr. A. De Paepe). The resulting construct was transformed into *Arabidopsis* ecotype Columbia, the *Arabidopsis cgl 1-2* mutant and the *Arabidopsis cgl 1-2* line homozygous for endoT (generation T3, line 5-3-1). The resulting T1 transformants were selfed to obtain twenty T2 segregating seed stocks per background. An acid buffer, that does not solubilizes the majority of the endogenous seed proteins, facilitates selective extraction of the ASP1 from ground seeds, yielding semi-pure extracts (see Material & methods). This allowed the T2 segregating seed stocks to be visually screened for ASP1 content by SDS-PAGE and Coomassie staining of the corresponding acid extracts (Supplemental Figure S1). Segregation analysis was subsequently performed to identify the single locus T-DNA integration lines and for each background, the line with the highest ASP1 accumulation and the ASP1 T-DNA integrated at a single locus and was retained and propagated to homozygosity.

Two patterns are observed when the recombinant ASP1 is produced in the different *Arabidopsis* backgrounds. In Columbia, two predominant, similarly stained bands are visible with a molecular weight just over 25 kDa and right below these, a third band of weak intensity (Figure 6). As ASP1 has two potential N-glycosylation sites, we hypothesize that the three bands represent different ASP1 glycoforms and thus correspond to ASP1 carrying no, one and two N-glycans. Apparently, most ASP1

is glycosylated, with an even distribution of the glycoforms with one and two N-glycans, whereas only little ASP1 carries no N-glycans. The pattern of ASP1 from Columbia is identical to that of the ASP1 expressed in the *cgl1-2* mutant and this seems to indicate that the presence or absence of complex plant glycosylation does not influence the production of ASP1 in *Arabidopsis* seed. When looking at the ASP1 derived from the GlycoDelete lines, a strikingly different pattern is visible (Figure 6). Here, only one band is present around 25 kDa and it coincides with the lowest band of the ASP1 from Columbia and the *cgl1-2* background. This confirms that the three bands of the Columbia and *cgl1-2* ASP1 indeed represent different glycoforms and that the lowest band corresponds to the unglycosylated ASP1. The more important conclusion to be drawn, however, is that the ASP1 seems fully susceptible to the endoT activity of the GlycoDelete line as no other glycoforms can be discerned on the gel

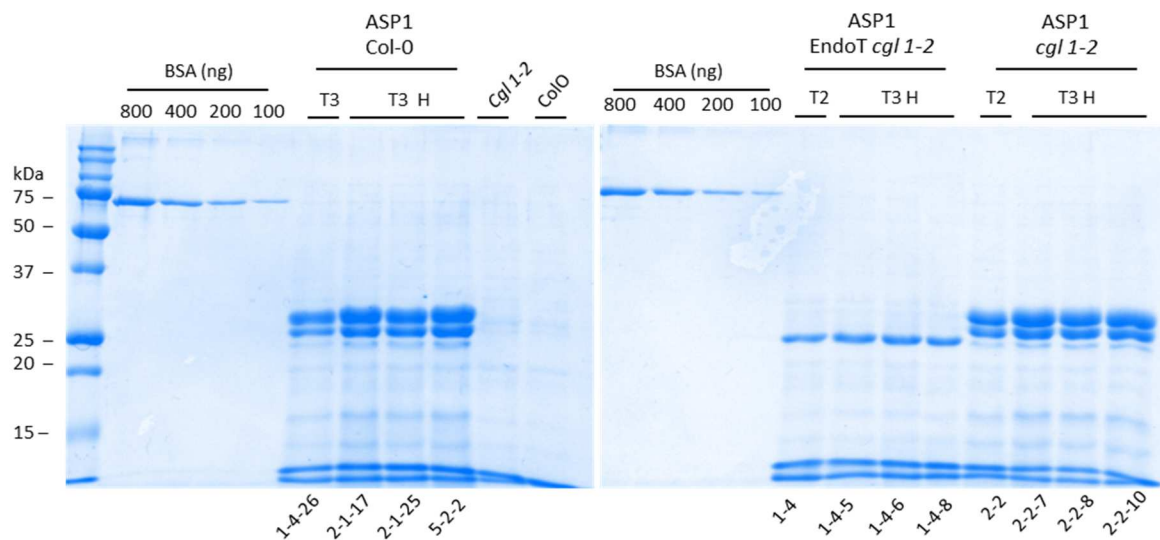


Figure 6. Analysis of recombinant ASP1 expression in different *Arabidopsis* backgrounds. Seed stocks of different lines producing recombinant ASP1 were extracted with an acid buffer at a fixed w/v ratio, separated by SDS-PAGE and Coomassie-stained. 10 μ l of extract was loaded and the gels were run under reducing conditions. A BSA standard was included to allow densitometric quantitation of the ASP1 present in the different extracts. Abbreviations used: BSA, bovine serum albumin; T2, single locus ASP1 T2 seed stock; T3, single locus ASP1 T3 seed stock; T3 H, homozygous ASP1 T3 seed stock; Col-0, *Arabidopsis* ecotype Columbia; *cgl1-2*, *Arabidopsis* complex-glycan-deficient 1-2 mutant. EndoT *cgl1-2*, *Arabidopsis* complex-glycan-deficient 1-2 mutant homozygous for endoT.

As the seed extracts are sufficiently pure, the signal intensity of the ASP1 bands was measured and used to determine the ASP1 content of the different extracts via the reference BSA standards that were also run on the gels. Together with the known w/v ratio at which the extracts were prepared, this allowed to calculate the ASP1 accumulation for each line, which was expressed as milligram

extractable ASP1 per gram of dry seed weight (Table 2). The accumulation of ASP1 in the Columbia and *cgl1-2* background, with an average of 27.8 and 27.2 mg per gram of dry seed weight respectively, is very similar for the three homozygous lines under study. This again supports the notion that the presence/absence of complex N-glycosylation does not alter the production of ASP1 in *Arabidopsis* seed. However, this picture is quite different for the GlycoDelete lines, where the average ASP1 accumulation for the three studied homozygous lines is only 7.1 mg per gram of dry seed weight. When compared to the Columbia and *cgl1-2* lines, this is approximately a four-fold reduction. Nevertheless, the ASP1 produced in the GlycoDelete background still reaches an accumulation of 3.5% of TSP, which is considerably higher than the 1% deemed the minimal threshold for economic viability of a PMP (Rybicki, 2009). It should be noted that for the *cgl* and the GlycoDelete ASP1 the three homozygous lines under study are derived from the same transformation event and hence, caution should be taken when interpreting the observed differences in ASP1 accumulation. However, lower accumulation of the ASP1 in the GlycoDelete background when compared to the *cgl* (and wild type) background seems a general tendency (compare Figure S1A and B). An additional detail revealed by table 2 is that a clear dosage effect can be observed when the seed stocks are propagated from heterozygous to homozygous and this characteristic is shared by all backgrounds.

Table 2. ASP1 content of different transformed lines.

Line	Background	Generation	mg/g seed	% of TSP
ASP1 1-4-26	Col-0	T3 He	20,4	10,2%
ASP1 2-1-17	Col-0	T3 Ho	28,6	14,3%
ASP1 2-1-25	Col-0	T3 Ho	24,8	12,4%
ASP1 5-2-2	Col-0	T3 Ho	29,9	14,9%
ASP1 1-4	<i>cgl1-2</i> endoT	T2 He	5,9	2,9%
ASP1 1-4-5	<i>cgl1-2</i> endoT	T3 Ho	8,0	4,0%
ASP1 1-4-6	<i>cgl1-2</i> endoT	T3 Ho	7,1	3,6%
ASP1 1-4-8	<i>cgl1-2</i> endoT	T3 Ho	6,0	3,0%
ASP1 2-2	<i>cgl1-2</i>	T2 He	19,8	9,9%
ASP1 2-2-7	<i>cgl1-2</i>	T3 Ho	28,7	14,4%
ASP1 2-2-8	<i>cgl1-2</i>	T3 Ho	23,5	11,8%
ASP1 2-2-10	<i>cgl1-2</i>	T3 Ho	29,3	14,7%

Table 2. Seed extracts were prepared from all the different recombinant ASP1 producing lines at a fixed w/v ratio and subjected to SDS-page. Densitometric analysis of Coomassie-stained gels, as shown in figure 5, allowed to calculate the ASP1 content of the different lines which was expressed as milligram ASP1 per gram of dry seed weight and also as % of TSP (for 200 mg TSP per gram of dry seed). Abbreviations used:

Col-0, *Arabidopsis thaliana* ecotype Columbia; *cgl1-2*, *Arabidopsis complex-glycan-deficient 1-2* mutant; *cgl1-2 endoT*, *Arabidopsis cgl1-2* mutant homozygous for *endoT*; He, heterozygous; Ho, homozygous; TSP, total soluble protein.

One possible explanation for the observed difference in accumulation of the ASP1 from the *cgl* background and the GlycoDelete line would be that deglycosylation of the ASP1 alters its solubility in the acid buffer used to extract the seeds. To check this the corresponding T3 seed stocks of both ASP1 variants were extracted three consecutive times with the acid buffer and a final fourth time with a buffer with a high concentration of detergent and reducing agent. The latter served to solubilize the pellet of seed debris and extract any remaining protein. All the collected extracts were analyzed by western blot to compare their ASP1 content and this shows that after three washing steps with the acid buffer a considerable fraction of both ASP1 variants remain in the pellet (Figure 7). However visual inspection of the western blot reveals no major difference in extraction pattern for both ASP1 variants and thus a difference in solubility can be discarded as an explanation for the earlier observed four-fold difference in accumulation of both variants. An interesting observation is made when looking at the pellet extraction of the GlycoDelete ASP1. Here, some faint bands of slightly higher molecular mass appear that are not present in the preceding acid extracts. These might correspond to intermediates of the ASP1 whose glycans have not yet been trimmed by *endoT*. This might represent the nascent ASP1 population in the ER.

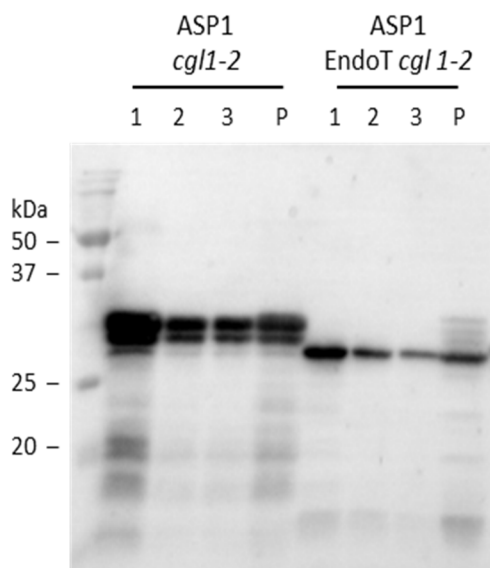


Figure 7. Solubility assay of the ASP1 produced in the *cgl1-2* and the GlycoDelete line. T3 seed stocks were extracted three times with an acid buffer (numbers 1 to 3) and a final fourth time with a buffer to solubilize the pellet (P). Extracts were prepared at a fixed w/v ratio and analyzed by western blot using a polyclonal rabbit α -ASP1 serum (1:1000). 10 μ l of extract was loaded and the gel was run under reducing conditions. Abbreviations used: *cgl1-2*, *Arabidopsis complex-glycan-deficient 1-2* mutant; *EndoT cgl1-2*, *Arabidopsis cgl1-2* line homozygous for *endoT*; P, pellet extraction; kDa, kilodalton.

EndoT accumulation level and stability over generations

We wanted to further characterize the seed-produced endoT by determining its accumulation level in both homozygous *cgl* lines and assessing its stability over multiple generations. To this goal, seed extracts were prepared from T3 seed stocks of both *cgl* lines homozygous for endoT, the endoT T3 2-2-7 *cgl1-1* and endoT T3 5-3-1 *cgl1-2* respectively, and from a T5 seed stock that is derived from the latter and co-expresses the ASP1 with the endoT. This line was created by supertransforming the parental T3 5-3-1 endoT *cgl1-2* line with the ASP1 construct and subsequent propagation to obtain the T5 5-3-1/1-4 endoT *cgl1-2* ASP1 seed stock that is homozygous for endoT and heterozygous for ASP1. The seed extracts were analyzed via ELISA and the obtained OD values plotted against a standard curve of purified fully processed endoT produced in the yeast *Pichia pastoris* to calculate the endoT content of the corresponding lines (Figure 8A). The seed and yeast-produced endoT were also subjected to western blot to assure recognition of both forms by the polyclonal α -endoT rabbit serum used in the ELISA and compare the patterns revealed for both enzymes (Figure 8C).

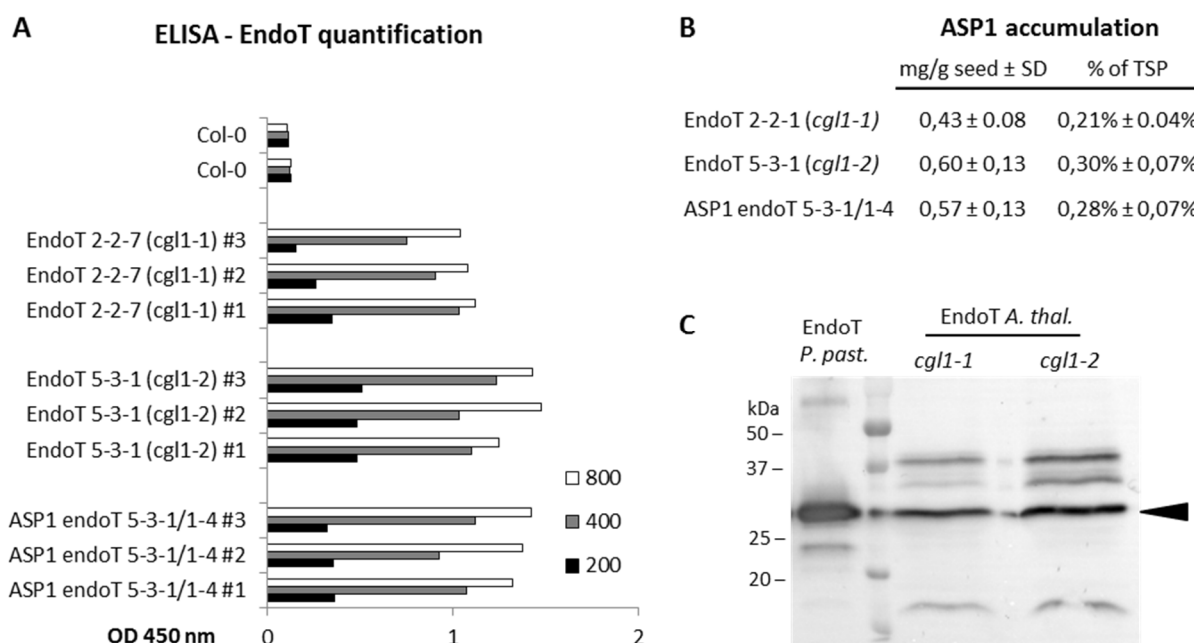


Figure 8. Quantification of seed-produced endoT and comparison to yeast-produced endoT. For each mutant background, *cgl1-1* and *cgl1-2*, the T3 seed stock of the line homozygous for endoT was selected to be analyzed (endoT 2-2-7 and 5-3-1, respectively), together with a T5 seed stock of the GlycoDelete line (endoT *cgl1-2*) that co-expresses endoT and ASP1 and is heterozygous for the latter (ASP1 endoT 5-3-1/1-4). Crude seed extracts were prepared and (A) subjected to ELISA using a polyclonal rabbit α -endoT serum. Are given, the dilution at which the extracts were coated and the obtained OD values. Numbers #1 to #3 represent biochemical repeats from the same seed stock. (B) The measured values were plotted

against an endoT standard curve and used to calculate the endoT content of the three analyzed lines. (C) Western blot comparison of the yeast- and seed-produced endoT detected by a polyclonal rabbit α -endoT serum (1:5000). Seed extracts were prepared at a fixed w/v ratio and 20 μ l was loaded, whereas for the purified *Pichia*-produced endoT 200 ng was loaded. The gel was run under reducing conditions and a black arrow indicates the endoT catalytic domain.

The ELISA shows that the endoT content of the GlycoDelete line in generation T3 (Figure 8A, EndoT *cg1-2*) and generation T5 co-expressing ASP1 (Figure 8A, ASP1 endoT *cg1-2*) is similar and, without even reaching 0.5% of TSP, quite low (Figure 8B). Supertransformation of the GlycoDelete line with the ASP1 construct under control of the same regulatory elements does not alter the endoT accumulation in the offspring line, which in generation T5 is still identical to that of the parental line (0.60 and 0.57mg, respectively) indicating stable inheritance of the endoT over several generations (T3 to T5). When the seed-produced endoT is compared to the yeast-produced endoT on western blot, the main signal of the catalytic domain of the endoT purified from yeast coincides with the major band around 33 kDa in the seed extracts, hereby supporting the earlier formulated hypothesis that this band indeed represents the fully processed endoT catalytic domain (Figure 8C).

Purification of recombinant ASP1

To obtain purified ASP1 for further in-depth analysis, the extraction of ASP1 by acid buffer from the seed was combined with a three-step purification protocol consisting of two consecutive ion-exchange chromatography steps followed by a final gel filtration step (Figure S2). When the protocol was applied to the wild type, *cg1*, and GlycoDelete ASP1, this yielded recombinant ASP1 of high purity (see Figure 9A and 9B). Under non-reducing conditions all the purified ASP1 variants are present as dimers that run at an apparent molecular weight of 37 to 50 kDa (Figure 9B). When the dimers are reduced the different ASP1 glycoforms can be discerned running at or right above 25 kDa (Figure 9A and 9B). It should be noted that prior to purification, the ASP1 proteins are predominantly present in the seed as dimers, but there is also a small fraction present as multimers and this is independent of the *Arabidopsis* background in which the ASP1 was produced (Figure S2). Nevertheless, as the ASP1 is encountered *in vivo* as a dimer of about 50 kDa (Meyvis et al., 2007), the ASP1 dimers were selectively separated from the ASP1 multimers during the final gel filtration step of the purification protocol and retained for further analysis. The presence of a fraction of the seed-produced ASP1s as multimers might point to a structural difference between the latter and the native ASP1 that occurs as dimers.

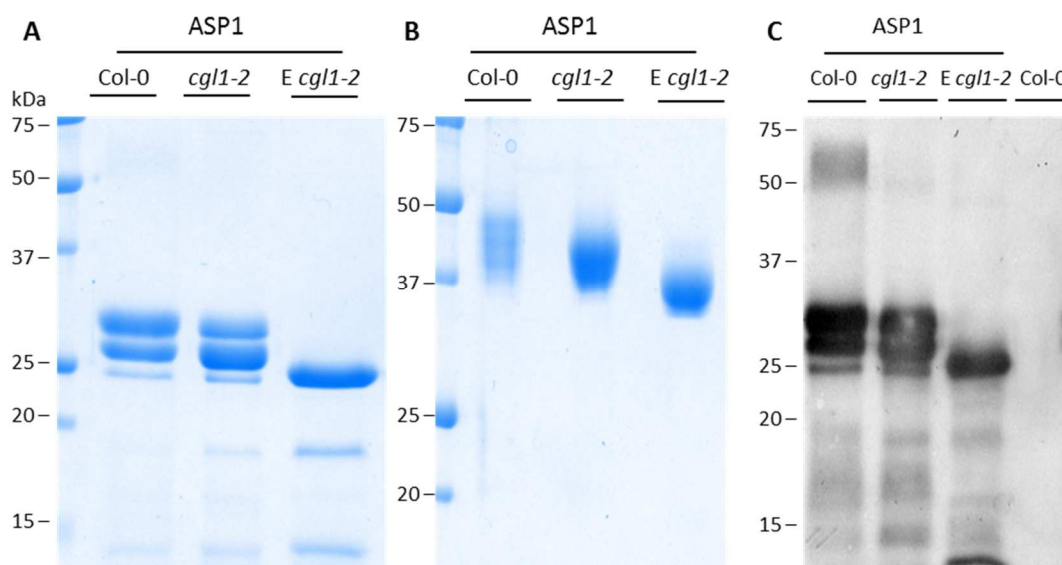


Figure 9. Characterization of purified recombinant ASP1. ASP1 produced in Col-0, *cgl1-2* and GlycoDelete seeds was purified by three consecutive chromatography steps and subjected to SDS-PAGE. Gels were Coomassie-stained and run under reducing (A) and non-reducing conditions (B) or analyzed via western blot using a polyclonal rabbit α -ASP serum (1:1000) and run under reducing conditions (C). 4 μ g of purified ASP1 was loaded onto the Coomassie-stained gels whereas 2 μ g was loaded onto the gel analyzed by western blot. Abbreviations used: Col-0, *Arabidopsis* ecotype Columbia; *cgl1-2*, *Arabidopsis* complex-glycan-deficient 1-2 mutant; E *cgl1-2*, *Arabidopsis* *cgl1-2* line homozygous for endoT.

Glycan analysis of ASP1 produced in wild type, cgl and GlycoDelete Arabidopsis seeds

ASP1 *cgl1-2*, but not ASP1 Col-0 is susceptible to *in vitro* endoT digestion

The profiles that the wild type and *cgl* ASP1 display under reducing conditions correspond to their different glycoforms, i.e. carrying no, one or two N-glycans (Figure 9A), and are almost identical pointing towards a similar N-glycan site occupancy in wild type and *cgl* *Arabidopsis*.

However, because of their production in the wild type and *cgl* *Arabidopsis*, the ASP1 variants should be decorated with complex and high-mannose N-glycans respectively, which cannot be distinguished as such on the gel. Nevertheless, the presence of these glycans on the recombinant ASP1 variants can be assayed by *in vitro* digest of both ASP1s by endoT, since it has substrate specificity for high-mannose, but not complex glycans. To this goal, purified wild type and *cgl* ASP1 was incubated with a dilution series of yeast-produced endoT subjected to SDS-PAGE and visualized via Coomassie staining (Figure 10).

As expected, the *cgl* ASP1 is susceptible to endoT treatment *in vitro* and the complete disappearance overnight of the bands representing glycosylated *cgl* ASP1 shows that it solely carries high-mannose N-glycans (Figure 10). On the other hand, the ASP1 wt is predominantly resistant to endoT digestion although the slight increase in intensity of the lower band indicates that some endoT sensitive glycans are also present. This points towards the acquisition of a majority of complex endoT resistant glycans together with some remaining high-mannose endoT sensitive glycans, which is in accordance with the glycan profile that is expected to be found on a secretory glycoprotein produced in wild type *Arabidopsis*.

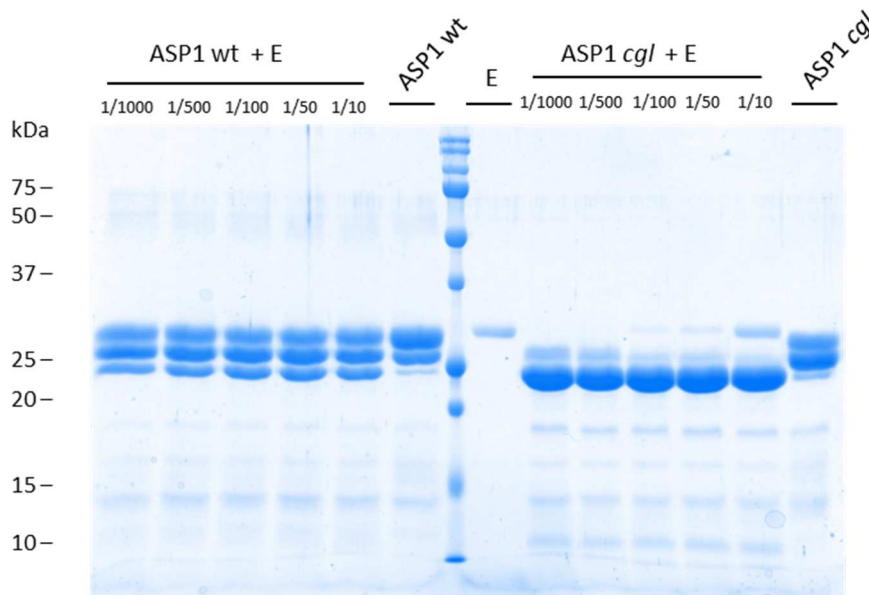


Figure 10. EndoT susceptibility of the ASP1 wt and *cgl*. Purified ASP1 wt and *cgl* was incubated overnight with a serial dilution of yeast-produced endoT (the ratio of endoT over target protein is indicated above the corresponding slots) and then subjected to SDS-PAGE and visualized by Coomassie staining. 5 µg of protein was loaded onto the gel that was run under reducing conditions. For each sample the ratio of endoT over the recombinant ASP1 is given. Abbreviations used: wt, *Arabidopsis* ecotype Columbia; *cgl*, *Arabidopsis* complex-glycan-deficient 1-2 mutant; E, endoT.

DSA-FACE analysis of recombinant ASP1 and the plant glycan substrate specificity of endoT

In concordance with the theoretically predicted profiles, the wt ASP1 carries endoT resistant complex glycans, the *cgl* ASP1 endoT sensitive high-mannose glycans and the GlycoDelete ASP1 glycans that are trimmed down to single GlcNAcs. To corroborate these findings and investigate the presence of

minor, previously undetected glycoforms, the ASP1s were further analyzed by DNA sequencer-assisted, fluorophore-assisted carbohydrate electrophoresis (DSA-FACE) and/or matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry.

The precise glycan profiles of the wt and *cgl* ASP1, predicted to contain predominantly complex glycans and high-mannose glycans respectively, were elucidated by releasing their glycans by PNGaseA digestion and sequential treatment of the total glycan pools with different glycosidases. DSA-FACE analysis of the glycan profile before and after each treatment then allowed annotation of each of the glycans in the total glycan pool (Figure 11B and 12E). If a glycan is susceptible to treatment with a specific glycosidase, its corresponding peak in the DSA-FACE electropherograms will shift after the reaction to that of the smaller product and hence comparison of both electropherograms can identify these glycans. After sequential treatments, eventually the exact structure of the glycans present in the electropherograms can be deduced (Laroy et al., 2006). Digestion with broad-spectrum mannosidase reveals glycans with terminal mannoses, whereas α 1,2-mannosidase identifies glycans with terminal α 1,2 mannoses only (Man₆ to Man₉). Treatment with hexosaminidase shows glycans with one or two terminal GlcNAcs (both complex and hybrid glycans) and any of these enzymes can be combined to digest structures with both terminal GlcNAcs and mannoses. To identify the peaks that represent complex α 1,3-fucosylated glycans the electropherogram of the total glycan pool released by PNGaseA can be compared to that of the glycans released by PNGaseF, which does not digest core α 1,3-fucosylated structures (Figure 12B and 12C).

When these techniques were applied to the wt ASP1 and its exact glycan profile was elucidated (Figure 11B, 12B and S3), several interesting observations were made. All glycans, except the high-mannose glycans that represent a small fraction of the total glycan pool, are core α 1,3-fucosylated and the majority of these also carry a bisecting β 1,2-xylose. Among these immunogenic plant glycans some paucimannosidic (Figure 11B peak 1 and 2) and hybrid structures (peak 6 and 7) can be discerned, but the predominant glycoforms are complex glycans with one terminal GlcNAc (and a fully processed α 1,3 arm, peak 2 and 3) or two terminal GlcNAcs (peak 4 and 5). This confirms the previous observation that wt ASP1 is mostly decorated with complex, endoT resistant, N-glycans. Indeed, all the glycans that are released by endoT treatment of wt ASP1 are exclusively of the high-mannose type (Figure 12D). Production of the ASP1 in the *Arabidopsis cgl* mutant drastically simplifies the glycan profile showing high-mannose glycans only with the structure Man₅ as the main glycoform (Figure 12E) and this is consistent with the lack of GnTI activity in the *cgl* background that blocks conversion of high-mannose to complex glycans.

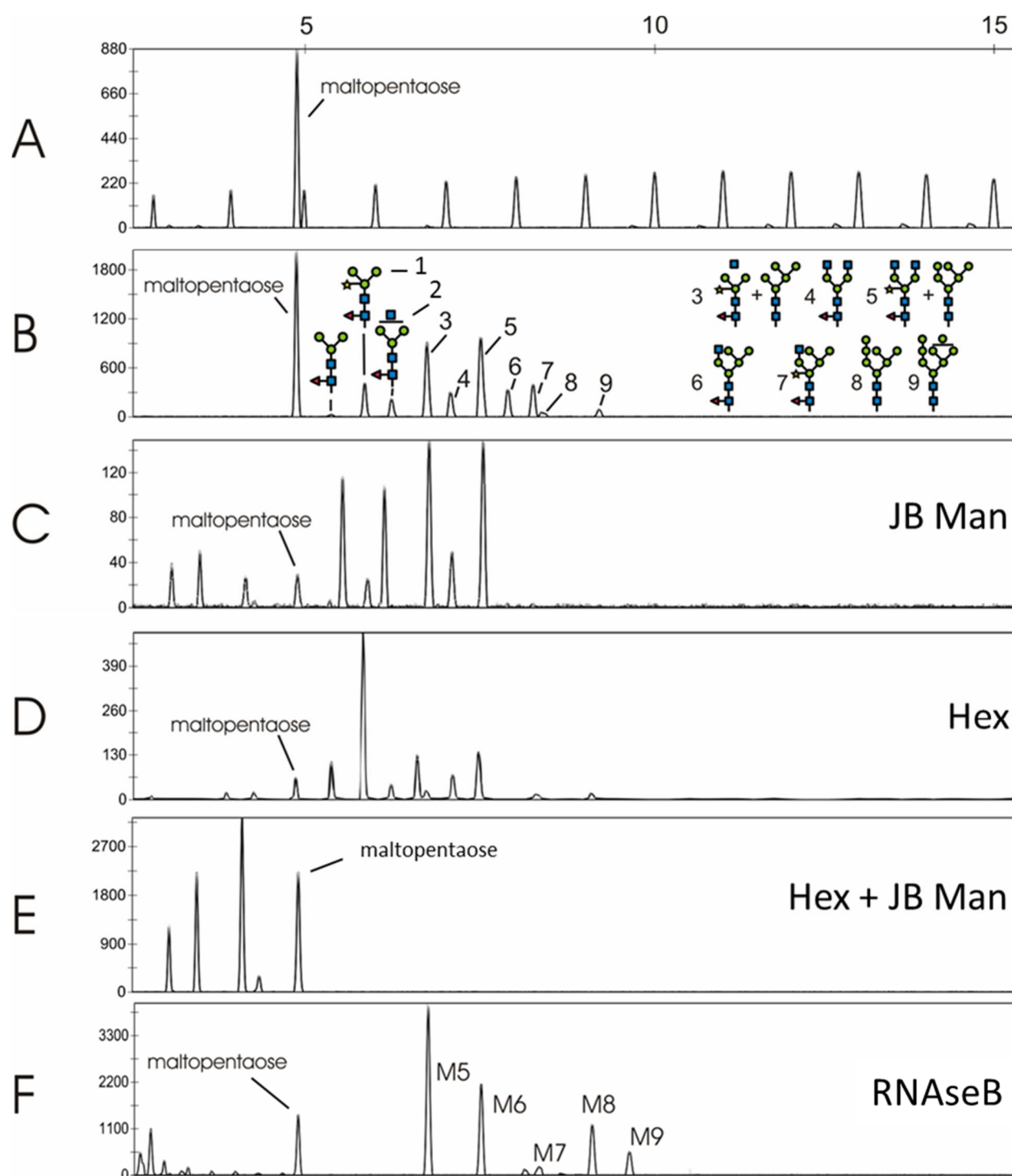


Figure 11. DSA-FACE analysis of the wt ASP1 glycan pool and glycosidase treatments thereof. The total glycan pool of wt ASP1 was released by PNGaseA digest, fluorescently labeled, treated with different glycosidases and subjected to DSA-FACE. (A) Oligomaltose standard (glucose polymer), the number of glucose units is indicated. (B) Total glycan pool released by PNGaseA and digested with (C) broad spectrum jack bean α 1,2,3,6-mannosidase, (D) hexosaminidase or (E) hexosaminidase and jack bean mannosidase. (F) High-mannose glycan profile of RNaseB. Maltopentaose was added to all samples as an internal standard. Abbreviations used: M5 to M9, Man₅ to Man₉; JB Man, jack bean α 1,2,3,6-mannosidase; Hex, jack bean β -N-acetylhexosaminidase.

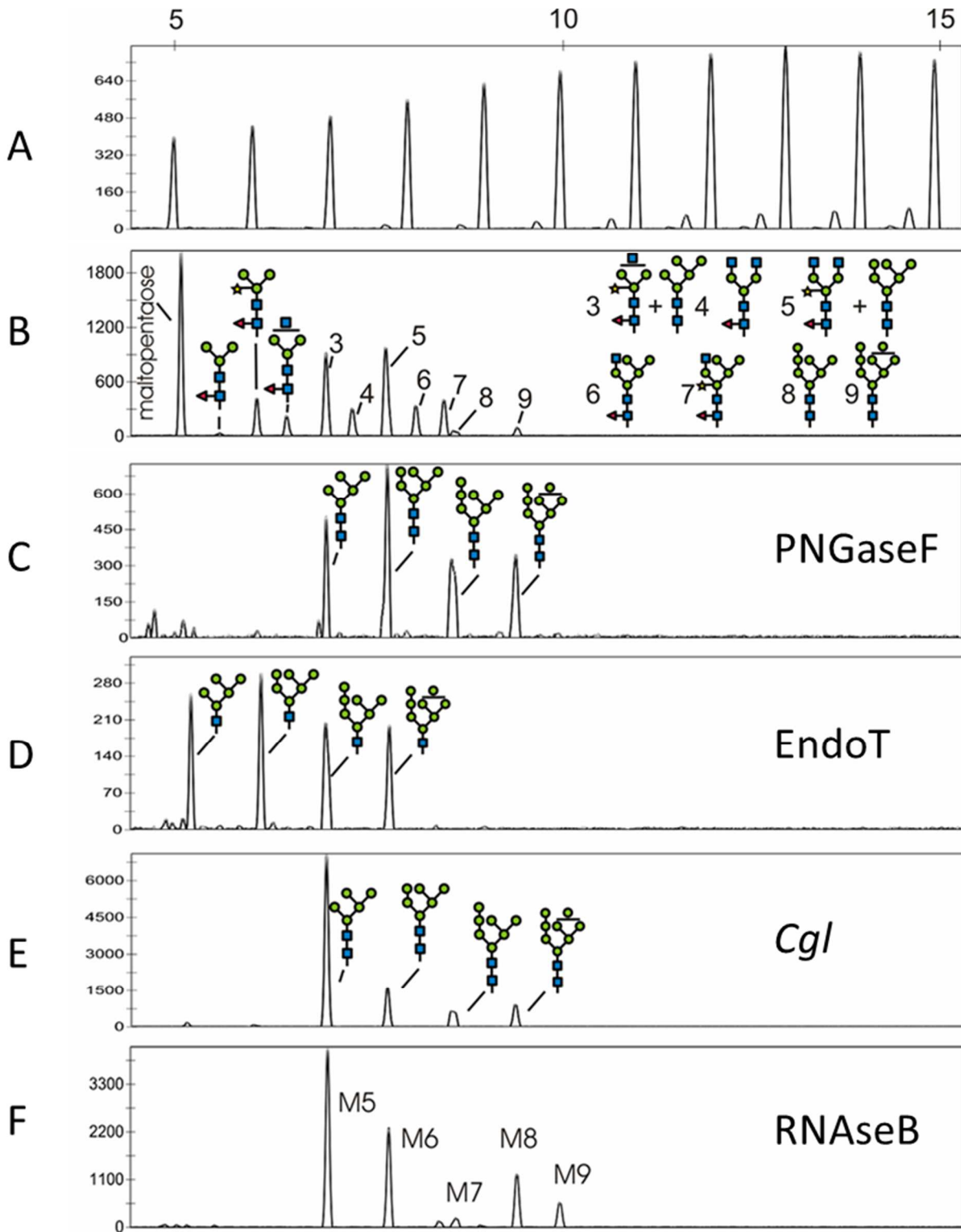


Figure 12. DSA-FACE analysis of the wt and *cgl* ASP1 glycan pool and the substrate specificity of endoT towards plant glycans. (A) Oligomaltose standard. (B) Glycans released from wt ASP1 by digestion with PNGaseA, (C) PNGaseF and (D) endoT. (E) *Cgl* ASP1 total glycan pool released by PNGaseA digest. (F) High-mannose glycan profile of RNaseB. Abbreviations used: M5 to M9, Man₅ to Man₉.

Mass spectrometry analysis of GlycoDelete ASP1

The only remaining glycan profile to be characterized in detail was that of the GlycoDelete ASP1, which could not be analyzed by DSA-FACE because of the small size of its trimmed glycans, which are not susceptible to PNGaseA or PNGaseF release. Hence, the GlycoDelete ASP1 was treated with the protease trypsin and the resulting tryptic peptides were subjected to MALDI-TOF spectrometry, which should allow precise identification of the glycan structures still present on the GlycoDelete ASP1.

Initially the purified GlycoDelete ASP1 was trypsinized in solution following a two-step protocol to ensure maximal proteolytic cleavage. The peptide mixture was then analyzed by MALDI-TOF spectrometry and the resulting spectrum screened for the two glycopeptides that were predicted to be released by tryptic digest of the GlycoDelete ASP1 (Figure 13A). All the peptides that could be identified in the spectrum are depicted with their corresponding apparent masses in Figure 13B and their peaks indicated on the spectrum (Figure 13C). Glycopeptide one and two are indeed present as glycoforms carrying a single GlcNAc and for both peptides tryptic variants can be found that are similarly decorated with a single GlcNAc. For glycopeptide one and its tryptic variant also the unglycosylated amino acid backbones can be discerned. Interestingly, no other glycoforms were present in the spectrum, pointing to complete digestion of the ASP1 glycans by endoT in the *Arabidopsis* GlycoDelete background. In general, several tryptic variants are visible, both for the glycopeptides as among the other peptides identified in the spectrum and these are the result of incomplete proteolytic digestion. Trypsin cleaves after a lysine (K) or arginine (R), however, sequences consisting of any two of the latter tend to be incompletely digested and apparently even the two-step digestion protocol could not ensure complete cleavage. Also some minor, unidentified peaks remain in the spectrum.

To further optimize the spectrum, the purified GlycoDelete ASP1 was run on a polyacrylamide gel prior to mass spectrometric analysis. The area corresponding to single GlcNAc and possible larger glycoforms was excised and the GlycoDelete ASP1 trypsinized in gel. The resulting peptides were then solubilized and subjected to MALDI-TOF mass spectrometry. When comparing the spectra obtained after both methods, an immediate and obvious difference is the reduction in number of tryptic variants and smaller peaks (Figure 13C and 14C). No tryptic variants of the glycopeptides are encountered and both glycopeptides are again present as glycoforms with single GlcNAcs although the peak corresponding to glycopeptide 1 is of low intensity (Figure 14C). No other glycoforms could be discerned confirming the previous observations that the ASP1 glycans are completely trimmed by the endoT in the GlycoDelete *Arabidopsis* background. By comparing the masses of the predicted tryptic peptides with those observed in the spectrum all major peaks could be annotated, except three peaks

at m/z 1975.02, 2185.15 and 2386.22. Nevertheless, aspecific cleavage of the GlycoDelete ASP1 *in silico* yields several peptides with m/z values that are nearly identical to those of the observed peak thus indicating that these probably also correspond to internal peptides of the GlycoDelete ASP1 (see Table S1).

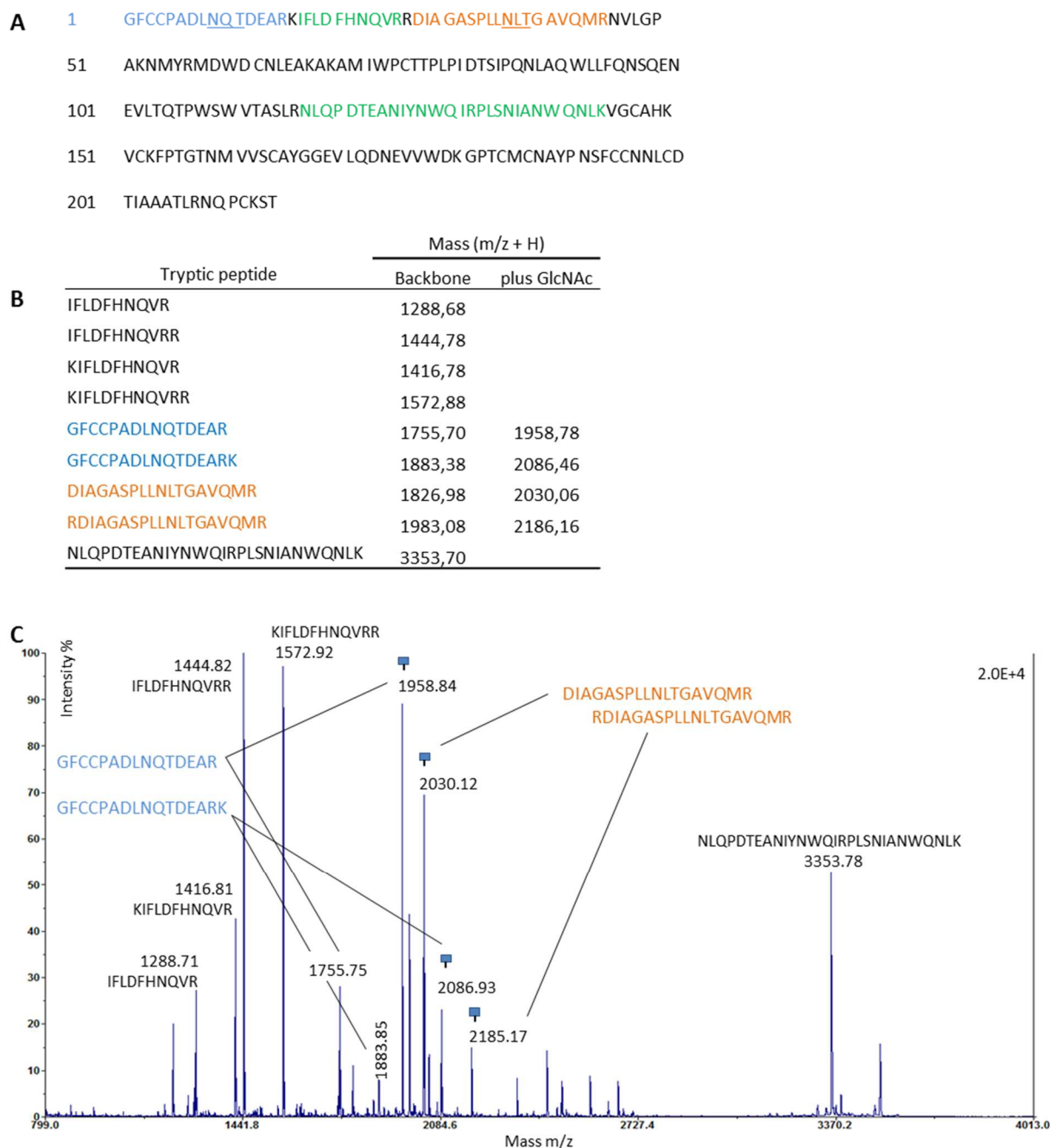


Figure 13. MALDI-TOF analysis of in-solution trypsinized glycodelete ASP1. Purified GlycoDelete ASP1 was carboxymethylated, trypsinized in-solution and subjected to MALDI-TOF spectrometry in the positive ion mode $[M+H]^+$. (A) The amino acid sequence of the GlycoDelete ASP1. Tryptic glycopeptide one and two are depicted in blue and orange, respectively, and the N-glycan attachment sites underlined. Other tryptic peptides found in the spectrum are indicated in green. (B) The different peptides identified in the spectrum and their masses (in kilodalton) given as the molecular weight plus one proton. (C) Mass spectrum of the trypsinized GlycoDelete ASP1. The peptide corresponding to a specific peak is indicated. A blue square represent a glycopeptide carrying a single GlcNAc.

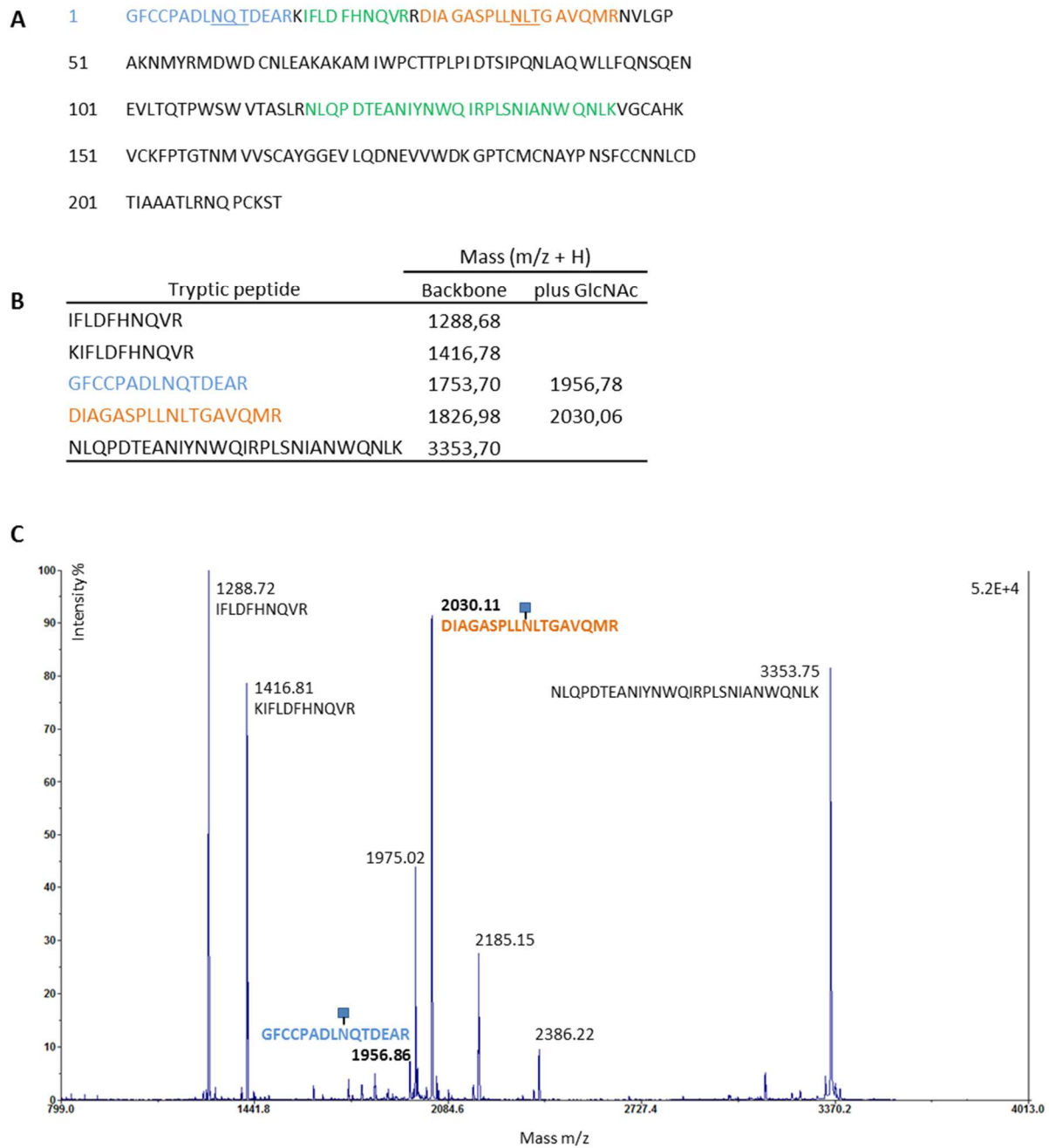


Figure 14. MALDI-TOF analysis of in-gel trypsinized glycodelete ASP1. Purified glycoDelete ASP1 was run on gel, carbamidomethylated and the corresponding band area excised and digested with trypsin. The resulting tryptic peptides were then solubilized and subjected to MALDI-TOF spectrometry in the positive ion mode $[M+H]^+$. (A) The amino acid sequence of the GlycoDelete ASP1. Tryptic glycopeptide one and two are depicted in blue and orange, respectively, and the N-glycan attachment sites underlined. Other tryptic peptides found in the spectrum are indicated in green. (B) The different peptides identified in the spectrum and their mass (in kilodalton) given as their molecular weight plus one proton. (C) Mass spectrum of the trypsinized GlycoDelete ASP1. The peptide corresponding to a specific peak is indicated. A blue square represent a glycopeptide carrying a single GlcNAc.

ASP1 carrying complex plant N-glycans, but not GlycoDelete ASP1 induces an anti-carbohydrate response in rabbits

Two groups of five rabbits were injected subcutaneously with GlycoDelete or wild type ASP1 following a prime-boost protocol. Rabbits received 35 µg of recombinant ASP1 mixed with complete freund's adjuvant at day zero and 35 µg mixed with incomplete freund's adjuvant at days 14, 28 and 56. Blood samples were taken on day zero prior to immunization (preimmune) and at days 28, 36 and 80 when the rabbits were euthanized. The goal of the experiment was to assess the immunogenicity of the wild type ASP1 with its fucose and xylose bearing complex plant glycans and compare this to the GlycoDelete ASP1 that carries truncated glycans.

After the final vaccination, all rabbits had developed high titers of IgG antibodies specific for the peptide backbone of the recombinant ASP1 (Figure 15A). These were measured by an direct ELISA in which denatured, reduced and deglycosylated ASP1 was coated as the antigen (Figure S4). Variation in titer was higher among the GlycoDelete ASP1 vaccinated group, but the average titer was similar and high for both groups (Figure 15A). To evaluate the immunogenicity of the glycan structures of both ASP1s, it was decided to study the presence/absence of IgG antibodies specific for plant glycans in the sera of the vaccinated rabbits. To measure the latter, a recombinant glycoprotein not related to ASP1 was used as a neutral carrier of glycan structures that served as antigens in ELISA to allow selective binding of glycan-specific IgGs. More specifically, a synthetic antibody was used that consisted of the antigen-binding domain of a camelid heavy chain only antibody (nanobody), i.e. the GFP-binding protein (GBP) coupled to the Fc domain of the murine IgG3 and this fusion protein is hereafter referred to as GBP-Fc. The recombinant GBP-Fc has two N-glycosylation sites located in its Fc-domain and had been expressed in different organisms, so different glycoforms were at hand. Production in *Pichia pastoris* yielded high-mannose glycoforms (Man₅ to Man₈) whereas the GBP-Fc produced in *Arabidopsis* seeds almost exclusively consisted of glycoforms with complex plant glycans bearing immunogenic fucose and xylose residues (PhD dissertation 2013, Dr. T. De Meyer; Figure 15C). It should be noted that approximately 30% of the *Arabidopsis*-produced GBP-Fc is unglycosylated (Dr. T. De Meyer, personal communication).

To probe the presence of plant-glycan specific IgGs, the sera of the vaccinated rabbits were subjected to ELISAs, using the high-mannose and immunogenic GBP-Fc glycoforms as coated antigens. Prior to the coating step, the purified glycoforms were reduced and denatured to expose their N-glycans that otherwise might have been shielded from the sera because of their inward orientation due to the fold of the Fc domain. The stripped GBP-Fc peptide backbone, obtained by reduction, denaturation and

digestion of the high-mannose glycoform with PNGaseF and monitored by mobility shift assay (Figure 15B), was also coated to measure non-glycan specific detection (background).

Three out of the five rabbits that had been injected with the recombinant ASP1 bearing complex plant N-glycans had generated IgGs specific for β 1,2-xylose and core α 1,3-fucose, confirming that these are indeed immunogenic (Figure 15A). Interestingly, the other two rabbits seemed unresponsive to the plant glycans despite having high titers of IgG specific for the ASP1 peptide backbone in their sera. No plant-glycan specific IgGs could be detected in the sera of the rabbits vaccinated with the GlycoDelete ASP1. For all ten rabbits, no signal could be measured when the high-mannose GBP-Fc was coated as antigen, pointing out that the detection of the fucose and/or xylose bearing GBP-Fc by three of the wild type ASP1 vaccinated rabbits was specific for the latter two residues.

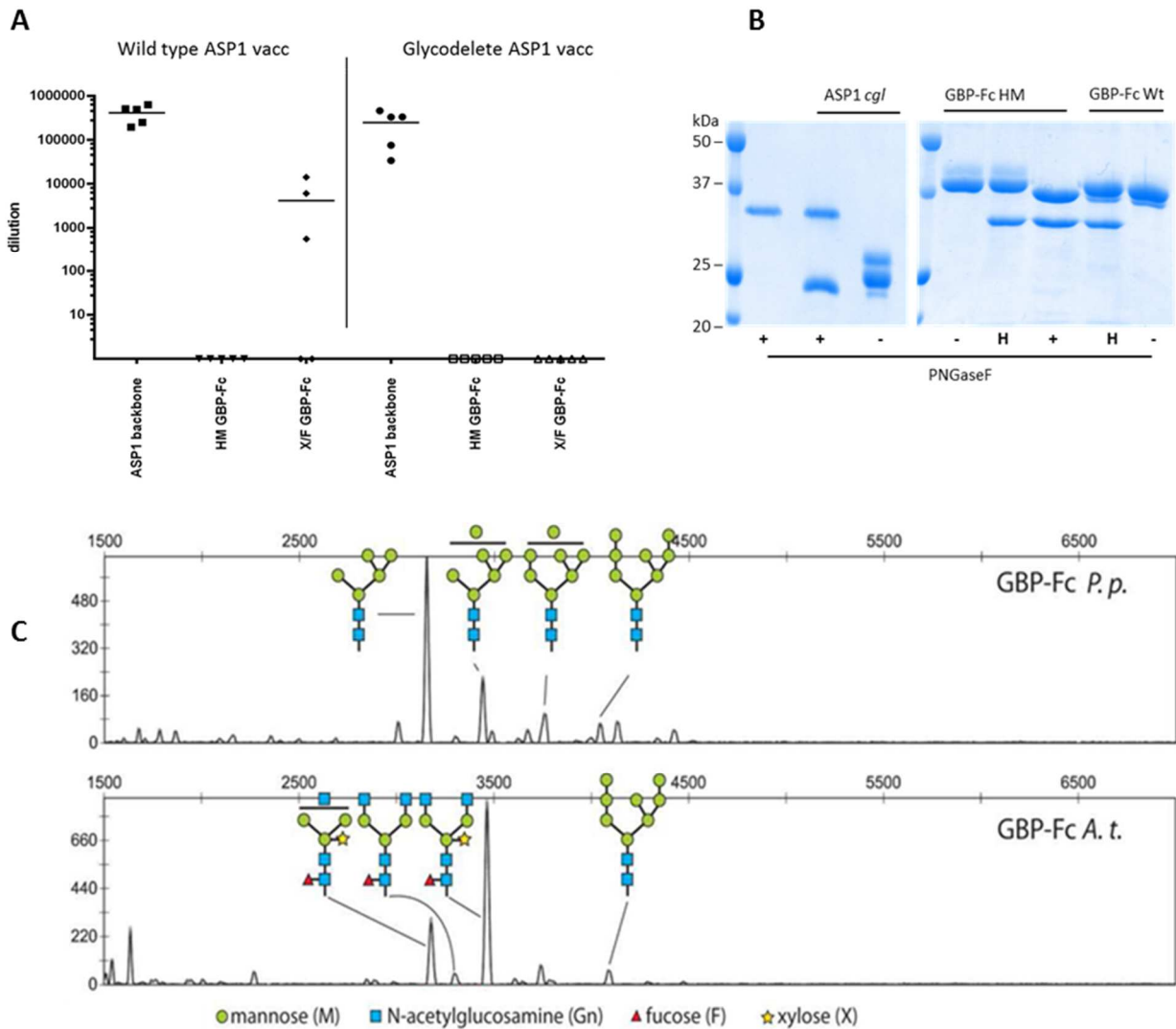


Figure 15. Overview of the ASP1 and plant-glycan specific IgG titers after vaccination together with the structure of the glycoforms of the GBP-Fc used during ELISAs. (A) Overview of the IgG titers measured in the sera obtained after the final bleeding of the rabbits. Both vaccinated groups are indicated above the figure. Are given, the IgG titers specific for the ASP1 peptide backbone and for both glycoforms of the GBP-Fc decorated with high-mannose and plant glycans respectively. The mean titer is depicted by a horizontal line. (B) Mobility shift assay to corroborate the deglycosylation of the ASP1 and GBP-Fc high-mannose glycoforms by PNGaseF digestion. (C) DSA-FACE electropherogram of the glycans released by PNGaseA from the GBP-Fc produced in yeast and *Arabidopsis* seed respectively. Abbreviations used: '–', no PNGaseF; '+', active PNGaseF; 'H', heat-inactivated PNGaseF; Vacc, vaccinated; kDa, kilodalton; *P.p.*, *Pichia pastoris*; *A.t.*, *Arabidopsis thaliana*; HM, high-mannose glycoform; X/F, complex glycoform bearing immunogenic xylose and fucose residues.

Discussion

The endoT was transformed into the seeds of the *Arabidopsis cgl* mutant and targeted to the *medial*-Golgi via the localization signal of the *Arabidopsis* β 1,2-xylosyltransferase. Expression analysis revealed that the majority of the endoT is both C and N-terminally processed, as occurs *in vivo*. Nonetheless, N-terminal processing would result in loss of the Golgi targeting signal as integral part of the cleaved propeptide and release the endoT from the *medial*-Golgi. It will be interesting to study the subcellular localization of the endoT in more detail to determine if the endoT, for example, accumulates in the apoplast, which would confirm its secretion and argue for loss of the *medial*-Golgi targeting signal. To characterize the engineered GlycoDelete platform, the endoT content of the GlycoDelete line was determined and shown to be stable over several generations and upon co-expression with a reporter glycoprotein. Despite the fact that the *β -phaseolin* promoter and *arcelin5-I* terminator, regulatory sequences derived from the common bean *Phaseolus vulgaris* and designed for high seed-specific expression of recombinant proteins, were used to drive endoT expression in the *cgl* seeds, endoT did not accumulate to more than 0.3% of TSP (or 60 milligram per gram of dry seed weight). A rather low number when compared to the levels, up to 15% and more of TSP, that have been achieved for synthetic antibodies produced in wild type *Arabidopsis* seeds using the same regulatory elements (De Buck et al., 2013; De Jaeger et al., 2002). Anyhow, massive endoT production during co-expression of a therapeutic protein is undesirable as it would deviate resources from the latter and might result being detrimental for its accumulation. Under the condition that sufficient endoT is present in the seed to completely process any therapeutic protein that is expressed in the GlycoDelete line, the low endoT content achieved actually represents an advantage rather than disadvantage and apparently, the low endoT content of the GlycoDelete line was perfectly capable of digesting a ten-fold excess of target protein (0.3 and 3.5% of TSP respectively), therefore demonstrating the functionality of the GlycoDelete platform.

A peculiarity here were the accumulation levels observed for the different recombinant proteins expressed in *Arabidopsis* seeds, with endoT showing the lowest number (0.3% of TSP), followed by the GlycoDelete ASP1 (3.5% of TSP) and similar and high levels for both the *cgl* and wild type ASP1 (13.6% and 13.9% of TSP respectively), despite sharing the same regulatory elements to drive their expression. It must be noted that, for cloning purposes, a shorter version (600 basepairs (bp)) of the *arcelin5-I* terminator was used to control endoT expression, whereas for the recombinant ASP1s the original terminator was employed (4000 bp). Although according to Goossens et al. (1999) the first 600 bp of the *arcelin5-I* terminator contain nearly all the necessary information for high expression, it was later shown that a transgene under control of the *β -phaseolin* promoter and the 600 bp *arcelin5-I* terminator did not achieve the same expression level as when the 4000 bp full-length terminator was

used, indicating that the shortened terminator is less potent (Dr. K. De Wilde, PhD dissertation 2013). Additionally, the lower endoT accumulation could also be a reflection of some of its intrinsic characteristic, e.g. the fact that it is a membrane-anchored protein which tends to be more difficult to express and it is probably the combination of these two factors that is responsible for the low endoT accumulation observed, which is actually beneficial for the GlycoDelete platform.

Another interesting feature was that the *Arabidopsis* line used for expression of the recombinant ASP1s seemed to influence the accumulation level with a four-fold reduction observed for the GlycoDelete ASP1 when compared to the *cgl* and wild type ASP1 (3.5% versus 13.6% and 13.9% of TSP respectively). The presence/absence of complex glycosylation did not seem to impact ASP1 expression as the *cgl* and wild type ASP1 accumulated to nearly identical levels, but co-expression with endoT apparently did, although it should be mentioned that the GlycoDelete ASP1 lines under study were offspring of the same transformation event and this should be kept in mind when extrapolating the observed differences between the GlycoDelete and the *cgl* and wild type ASP1. Nevertheless, comparison of supplemental figure S1A and S1B seems to indicate a general tendency of the recombinant ASP1 to accumulate less in the GlycoDelete background. As the only difference between the GlycoDelete and *cgl* ASP1 resides in the presence/absence of the endoT gene construct and the corresponding translated protein in its host, it is plausible to attribute the lower accumulation of the GlycoDelete ASP1 to one of the latter two and several explanations are possible. Firstly, despite leaving the proximal GlcNAc attached to the peptide backbone, the trimmed ASP1 might just be less stable and hence, not accumulate to the same levels as its *cgl* and wild type counterparts. Secondly, although the endoT is targeted to the *medial*-Golgi to avoid interference with protein folding in the ER, the GlycoDelete ASP1 and endoT are under control of the same promoter and thus have an identical temporal expression pattern what allows for contact between both proteins in the ER and this might interfere with the correct folding of a fraction of the ASP1 leading to its degradation by the ER-associated degradation pathway (ERAD). Thirdly, since both proteins share the same promoter they might compete for resources during production lowering final GlycoDelete ASP1 accumulation, although this seems little credible given that the endoT accumulation in the GlycoDelete line is low (0.3% of TSP) and was identical before and after co-expression of the ASP1.

A final possibility that cannot be excluded, is that by mere chance only low GlycoDelete ASP1 accumulators were found among the 20 transformants that were screened (supplemental Figure S1A) and that by amplifying the initial screen, transformants would have been picked up with accumulation levels similar to those observed for the *cgl* and wild type ASP1. To exclude this chance factor, a well characterized *cgl* ASP1 producing line should be supertransformed with the endoT construct, the opposite operation of how the GlycoDelete ASP1 was generated during this work, and the

accumulation of the GlycoDelete ASP1 in the offspring should be compared to the parental line without endoT. This will allow an exact determination of the influence of the endoT on the ASP1 accumulation and unveil the mechanism for lowering GlycoDelete ASP1 accumulation.

With 3.5, 13.6 and 13.9% of TSP the accumulation levels achieved for the GlycoDelete, *cgl* and wild type ASP1, respectively, compare favorably to those obtained for other recombinant proteins that were produced in *Arabidopsis* seeds using the same expression cassette and displayed accumulation levels generally ranging from <0.1 up to 15% of TSP (Loos et al., 2011a; Loos et al., 2011b; Morandini et al., 2011; Van Droogenbroeck et al., 2007). Apparently, the *Arabidopsis* seed platform seems to be exceptionally suited for the production of antibodies and derivatives thereof, given that these make up the higher end of the accumulation range with even an all-time record yield of 36.5% of TSP for a murine single-chain variable fragment (De Buck et al., 2013; De Jaeger et al., 2002; Viridi et al., 2013). On the lower end of the range, enzymes and antigenic proteins are found and these often need extensive optimization to reach reasonable yields (Morandini et al., 2011; Piron et al., 2014). Despite the fact that the *cgl* and wild type ASP1 are not structurally related to antibodies, they also rank among the high accumulating recombinant proteins produced in *Arabidopsis* seeds. In fact, the 13.6% of TSP reached by the *cgl* ASP1 is to our best knowledge the highest yield ever reported for a recombinant protein expressed in this specific *Arabidopsis* background. The seeds of the *cgl* mutant have been regularly used as a platform for the production of enzymes destined for enzyme replacement therapy of lysosomal storage diseases because of the high-mannose glycans introduced on the enzymes in this background that allow them to be specifically targeted to macrophages (Downing et al., 2006; Downing et al., 2007; He et al., 2012a; He et al., 2012b). Several enzyme variants have been produced in the *cgl* seeds and despite extensive optimization and the use of similar regulatory elements to drive expression, i.e. the *arcelin5-l* promoter and terminator, the highest yield reported for a recombinant variant of human α -L-iduronidase was only 5.7% of TSP (He et al., 2013). In summary, recombinant ASP1 accumulated to high levels in both wild type and *cgl* *Arabidopsis* seeds and although the accumulation of the GlycoDelete ASP1 was fourfold less, it still by far exceeded 1% of TSP, a value that has been designated to be the minimum required for economic viability of a PMP (Rybicki, 2009).

To study the immunogenic potential of the complex plant glycans present on the wild type ASP1 but absent on the GlycoDelete ASP1, rabbits were immunized with purified GlycoDelete and wild type ASP1. More specifically, this experiment was performed to determine if the plant specific β 1,2-xylose and core α 1,3-fucose that were almost ubiquitously present on the glycans of the wild type ASP1, had induced glycan-specific antibodies and in case this was true, to see if this response was lacking among the GlycoDelete ASP1 immunized rabbits. Administration of the GlycoDelete and wild type ASP1

resulted in high titers of IgG antibodies specific for the ASP1 peptide backbone in all immunized animals and indeed, also three out of five rabbits that had received the wild type ASP1 had generated IgG antibodies against the glycans carried by the wild type ASP1. The specificity of these antibodies for the plant specific β 1,2-xylose and core α 1,3-fucose was demonstrated by the fact that only complex plant glycans, but not high-mannose glycans were recognized by the sera with glycan specific IgGs. Curiously, the remaining two rabbits from the wild type ASP1 immunized group had not mounted a glycan specific antibody response and this probably reflects genetic differences in susceptibility to these potential immunogenic plant glycans. As was expected, no glycan specific IgG antibodies could be detected in the sera from the rabbits immunized with the GlycoDelete ASP1 confirming the lack of immunogenic glycans on the latter. To sum up, this trial once more validates the immunogenic nature of the structures β 1,2-xylose and α 1,3-fucose that are omnipresent on complex plant glycans but absent on mammalian glycans (reviewed by Altmann (2007)). However it also identifies genetic disposition as a key factor in determining the individual outcome of exposure to a potentially immunogenic structures, i.e. if it is tolerated or if an antibody response is mounted against it. Most importantly, It also demonstrates that the GlycoDelete platform can be effectively used for the production of therapeutic proteins decorated with minimal, homogeneous glycans that are devoid of β 1,2-xylose and α 1,3-fucose, structures whose presence could otherwise invoke an anti-carbohydrate response upon parenteral administration of the therapeutic and could be potentially dangerous.

The GlycoDelete platform presented here is a very ingenious approach to tackle the problem that immunogenic plant glycans (and glycan heterogeneity) pose. It does not suffer from the drawbacks that other strategies aimed at 'humanizing' plant glycans display. It is based upon a single inactivated endogenous glycosyltransferase and one introduced glycosidase, whereas other methods require the knock-out and knock-in of multiple genes and these manipulations are in general not completely efficient, which can actually even result in increased glycan heterogeneity. This is in sharp contrast with the GlycoDelete seed platform that fully trimmed the glycans on a ten-fold excess of ASP1 to homogeneity. None of the strategies for which proof-of-concept was delivered on a lab scale has so far been introduced in a commercial crop and industrialized, because of the excessively large timescale it would take to introduce all these traits to homozygosity, and then into an elite cultivar. In soybean, for example, a year and half is needed between transformation and the generation of seeds homozygous for the introduced trait (Dr. V. Viridi, personal communication). Stable integration of the six genes needed for sialylation of plant glycans would then take about ten years, without even having dealt with the introduced glycan heterogeneity and the inactivation of the genes responsible for the immunogenic glycans, and this is where the GlycoDelete technology could mark a difference. Its two-component setup would greatly reduce development time and facilitate its transfer to any crop

for which a GnTI knock-out can be obtained. A task that is considerably less cumbersome now meganucleases and genome-editing technologies are making their entry into research and development. The GlycoDelete technology has great potential and could be put to use in a whole range of applications. Actually, any protein that is produced in plants, but should lack immunogenic plant glycans and show great homogeneity, would benefit from the latter. One could envisage the large scale production in soybean or pea, whose seeds have a remarkably high protein content, of monoclonal antibodies (or Fc-fusion proteins) for which effector functions are unwanted, or vaccines that show enhanced efficacy when minimally glycosylated. Needless to say, the options for the use of the GlycoDelete technology are ample and it will be interesting to see the future development of the technology *in planta*.

Immunization of calves with native ASP1 provides sufficient protection to allow them to build up natural immunity against *Ostertagia*. However, because of difficulties to obtain large quantities of native ASP1, there is great interest to produce recombinant ASP1s and despite the setbacks experienced with the recombinant ASP1s from baculovirus-infected insect cells, *Pichia* and wild type *Arabidopsis*, i.e. their non-protective nature, hopes are still high that at some point recombinant ASP1 can be produced for vaccine purposes. To achieve the latter, the structural element that imbows the native ASP1, but not the recombinant versions, with protective capacities must be identified and analyses are ongoing to solve this issue. Although the glycans present on the native ASP1 did not influence the binding of neutralizing antibodies, it will be interesting to analyze the effect of the different glycans present on the wild type, *cgI* and GlycoDelete ASP1 upon immunization and this will contribute to a better understanding of the characteristics of the ASP1, which will eventually lead to the elucidation of the enigmatic 'protective element'.

Material and methods

Cloning and generation of transgenic lines

The coding sequences of the *Trichoderma reesei* endoT and the *Ostertagia ostertagi* ASP1 were derived from those deposited at GenBank under accession number [4G2U_A](#) and [CAZ16624.1](#), respectively. The recombinant endoT and ASP1 sequences that were cloned as described below and used to transform various *Arabidopsis* lines are given in the supplemental information (SI1 and SI2).

The sequence coding for the endoT without signal peptide was extended *in silico* at its 5'-end with the Kozak motif (CCACC) followed by a sequence encoding the first 32 amino acids that make up the medial-Golgi targeting signal of the *Arabidopsis* β 1,2-xylosyltransferase (Nagels et al., 2011b; Pagny et al., 2003). The resulting gene was optimized to resemble the codon usage of endogenous *Arabidopsis* seed storage proteins and chemically synthesized flanked by attL1-L2 sites. A multisite gateway recombination reaction (all GatewayTM reactions were performed according to the manufacturer's protocol) transferred the gene to a destination vector that contains a T-DNA segment carrying the *bar* gene for phosphinothricin resistance and a seed-specific expression cassette under control of the *β -phaseolin* promoter and the *arcelin5-I* terminator.

The coding sequence of the ASP1 was extended *in silico* at its 5'-end with the Kozak motif (CCACC) plus a sequence encoding the ER-targeting peptide of the *Arabidopsis* 2S albumin seed storage protein (Krebbers et al., 1988). The resulting gene was codon optimized, flanked by attL1-L2 sites and chemically synthesized. An LR recombination reaction transferred the gene to the gateway compatible pPphasGW destination vector that contains a T-DNA segment carrying the *nptII* gene for kanamycin resistance and a seed-specific expression cassette under control of the *β -phaseolin* promoter and the *arcelin5-I* terminator (Morandini et al., 2011).

All the genes were sequence-verified prior to transfer to the destination vectors. Heat shock then delivered all expression vectors to *Agrobacterium* strain C58C1RifR containing the pMP90 virulence plasmid (Koncz & Schell, 1986) and the resulting strains were used to transform *Arabidopsis thaliana* (L.) Heyhn, ecotype Columbia 0 and the allelic *complex-glycan deficient 1-1* and *1-2* mutants, by floral dip (De Buck et al., 2012; von Schaewen et al., 1993). The *Arabidopsis thaliana* *cgl1-1* and *1-2* mutants, stock number N6192 and N16366 respectively, were obtained from the European *Arabidopsis* Stock Centre (Nottingham, UK). Transformed seeds containing the T-DNA were identified and lines propagated to homozygosity as detailed in (De Buck et al., 2012). In short, T1 seeds obtained after floral dip were grown on selective Murashige and Skoog medium for three to four weeks at 21 C° on a 16-h light / 8-h dark cycle. Twenty transformants were then transferred to soil and grown under the

same conditions. These were selfed and when seed setting was complete, transferred to a room at 25 °C and low humidity until completely dry. The dry seeds of each transformant were harvested to obtain twenty T2 seed stocks. These stocks were screened for the presence of recombinant protein by western blot analysis (for endoT) or Coomassie staining of polyacrylamide gels (for ASP1) as detailed below and seven stocks with high accumulation of the recombinant protein were analyzed for the number of integrated T-DNA loci. As the T1 transformants are per definition hemizygous, the T-DNA locus number can be determined by the segregation ratio in the T2 generation. To identify the single locus T-DNA integration lines, seeds of segregating T2 stocks were sown on selective medium and the observed ratio of germinated over sensitive seeds was compared to the expected 3:1 segregation ratio for single locus lines using a χ^2 statistical test (De Buck et al., 2012). For each antigen, the T2 stock with the highest antigen content and the T-DNA integrated at a single locus in the genome was retained and propagated to homozygosity.

Protein extraction, protein concentration determination, SDS-PAGE, Coomassie staining and western blot analysis

To characterize the endoT accumulation in transgenic seed, TSP was extracted as described elsewhere (De Buck et al., 2012) with minor modifications: 1 ml of cold extraction buffer (50 mM Phosphate buffer, pH 7.8, 0.3 M NaCl and supplemented with an EDTA-free, proteinase inhibitor cocktail (1 tablet per 50 ml; Roche)) with or without the detergent n-Dodecyl β -D-maltoside (1% w/v) was added to 10 mg of liquid nitrogen (N2) frozen, crushed seeds. Samples were vortexed for 1 minute and centrifuged at approximately 20,000 g and 4°C for 30 minutes. 700 μ l supernatant was removed, snap-frozen in liquid N2 and stored at -20°C. Total protein concentration was determined by the Lowry method as described in De Buck et al. (2012). The ASP1 was extracted from the seed following the same protocol that was applied to the endoT except that an acid extraction buffer was used (0.1 M sodium acetate at pH 4.4 supplemented with CHAPS (0.1% w/v) and an EDTA-free, proteinase inhibitor cocktail (1 tablet per 50 ml; Roche)). During the solubility assay, ASP1 was extracted from the seeds three times with 0.1 M sodium acetate buffer at pH 4.4 and a final fourth time using the same buffer but supplemented with 5% v/v β -mercaptoethanol and 8 M urea.

Seed extracts containing TSP were separated by reducing SDS-PAGE. To visualize proteins via Coomassie staining, the gels were submerged in InstantBlue™ (Expedeon) solution for an hour and then washed three times for 20 minutes with double distilled water (ddH2O). Stained gels were photographed with the Molecular Imager® ChemiDoc™ XRS+ Imaging System (Biorad) and densitometric analysis of the visualized proteins was performed with the Image Lab 3.0 software

(Biorad). For western blot analysis the proteins were transferred electrophoretically from a polyacrylamide gel onto Immobilon-P polyvinylidene fluoride membrane (Millipore, Billerica, MA). Blots were blocked in 4% (w/v) skimmed milk and 0.1% (v/v) Tween 20 in PBS. Polyclonal rabbit α -endoT and α -ASP1 sera, diluted 1:10.000 and 1:5000 in blocking buffer, respectively, were used to detect the recombinant proteins in combination with an anti-rabbit IgG HRP-linked whole antibody diluted 1:5000 in blocking buffer (from sheep, GE healthcare). Western blotting was further performed as described in De Buck et al. (2012) and the released chemiluminescent signal measured with the abovementioned imaging system. Quantification of signal intensities was performed with the Image Lab 3.0 software (Biorad). To identify high-mannose bearing glycoproteins western blots were developed using ConA coupled to HRP: after transfer of the SDS-PAGE separated proteins onto the blotting membrane, remaining free sites were blocked with 3% BSA and 0.05% (v/v) Tween 20 in PBS and the membrane incubated with ConA-HRP (1 mg/ml in blockbuffer) with or without the ligand methyl- α -D-mannopyranoside (100 mM, in blockbuffer). Western blotting was further performed as described above.

PNGaseF digestion of recombinant endoT

Seed extracts containing TSP were treated with peptide N-glycosidase F (PNGaseF, New England Biolabs) to analyze the endoT glycosylation status. Glycosidase reactions were performed according to the manufacturer's protocol. In short, to 9 μ l of extract, 1 μ l of 10x denaturing buffer was added and the sample was incubated at 100°C for 10 minutes. Then, 2 μ l G7-buffer, 2 μ l 10% NP-40 and 2 μ l PNGaseF was added to the sample. Sample volume was adjusted to 20 μ l with ddH₂O, after which the sample was incubated at 37°C for one hour and analyzed by western blotting.

ELISA to determine endoT accumulation level

The endoT content of specific seed stocks was analyzed by ELISA. Seed extract were prepared in binding buffer with detergent as detailed above, dilution series were made in duplicate with sodium carbonate as the buffer (50mM NaHCO₃, pH 9.6), loaded onto Maxisorp 96-well plates (Nunc, Sigma Aldrich) and coated overnight at 4°C. Plates were washed three times with Tween (0.1% v/v) in PBS between all following steps unless stated differently. PBS with BSA (1% w/v) and Tween (0.1% v/v) was used to block the wells after which 100 μ l polyclonal rabbit α -endoT was added (diluted 1:2.500 in blocking buffer). Subsequently 100 μ l donkey α -rabbit IgG-HRP (Sigma Aldrich) was added as secondary antibody (diluted 1:5000 in blocking buffer) and the wells were washed five times.

Incubation with 100 µl TMB (Sigma Aldrich) allowed color development that was stopped after 30 minutes by addition of 100 µl 1M HCl after which the plates were read at 450 nm in a VERSAmax tunable microplate reader. Measured OD values were plotted against a standard curve derived from duplicate dilution series of purified endoT and used to calculate the endoT content of the different stocks.

Recombinant ASP1 purification

The acid extraction of the recombinant ASP1s from the seeds was coupled to a three-step purification protocol consisting of cation and anion exchange chromatography followed by gel filtration. Seed extracts were prepared in binding buffer (40mM sodium acetate, pH 4.4, supplemented with an EDTA free proteinase inhibitor cocktail (1 tablet per 50 ml; Roche)) by adding frozen and pulverized seeds to cold buffer at a fixed w/v of 1 g seeds per 70 ml buffer. The extract was vortexed and centrifuged for 30 minutes at 4°C and 40,000 g to remove cell debris. The supernatant was removed, passed through a GF-prefilter (Sartorius) and a Millex®-HP filter unit (Millipore) and then flowed over the cation exchange column (Hitrap SP HP 1 ml; GE Healthcare) fitted onto an Äkta Explorer® purification system. Bound ASP1 was eluted by passing a gradient of elution buffer (40 mM sodium acetate, pH 4.4, 1M NaCl) over the column. Subsequently the eluted fractions containing the majority of the ASP1 were pooled, diluted in 20mM TrisHCl to lower the conductivity, adjusted to pH 8.0 and flowed over the anion exchange column (Hitrap Q HP 1 ml; GE healthcare) fitted onto the same purification system. Bound ASP1 was released from the column by a gradient of elution buffer (20 mM TrisHCl, pH 8.0, 1 M NaCl). Again, the eluted fractions containing most ASP1 were retained, pooled and subjected to a final gel filtration step by passage over a HiLoad Superdex 200 pg column (GE healthcare), with LPS-free PBS as the mobile phase, fitted onto the same purification system. Removal of impurities and/or degradation products was assayed by Coomassie staining of polyacrylamide gels and whenever necessary, pure, gel filtrated ASP1 was concentrated using Amicon® Ultra centrifugal filters (Millipore, 3 kDa molecular weight cutoff). The ASP1 concentration was derived from the sample optical density at 280 nm measured by a NanoDrop® ND-1000 spectrophotometer. Samples were snap-frozen in liquid N₂ and stocked at -20°C.

Recombinant ASP1 glycan analysis by endoT digest

To check the susceptibility of the glycan structures carried by wild type and *cg1* ASP1 to digestion with endoT, purified ASP1 (in PBS) was incubated overnight at room temperature with serial dilutions of

active, purified endoT produced in the yeast *Pichia pastoris*. Sample glycosylation status was subsequently checked by a mobility shift assay on Coomassie stained polyacrylamide gels.

Recombinant ASP1 glycan analysis by glycosidase treatment and DSA-FACE

To identify the glycan structures present on wild type and *cgl* ASP1, purified recombinant ASP1 was digested with the endoglycosidases PNGaseA and PNGaseF to release its glycan structures that were then fluorescently labeled and analyzed by DSA-FACE via the plate method as described in (Laroy et al., 2006). In short, reduced and denatured, and for PNGaseA digest also trypsinized, ASP1 was bound on polyvinylidene fluoride membranes in a 96-well plate. The washed, carboxymethylated, protein was digested with PNGaseF or PNGaseA, the supernatants transferred to a reaction tube and evaporated. Glycans were fluorescently labeled with 8-amino-1,3,6-pyrenetrisulfonic acid (APTS) and gel filtrated by passage over G10 Sephadex in 96-well plates to desalt and remove excess label. After evaporation and resuspension in nanopure water, the glycans were eventually analyzed on a DNA-sequencer. Sequential treatment and analysis of the released glycans with different exoglycosidases allowed elucidation of the glycan structures present on the wild type and *cgl* ASP1.

Recombinant ASP1 glycan analysis by MALDI-TOF spectrometry

The truncation of the glycans on the ASP1 produced in the GlycoDelete *Arabidopsis* was analyzed by MALDI-TOF mass spectrometry. Trypsinization of the GlycoDelete ASP1 prior to mass spectrometry occurred in-solution or in-gel according to the protocols described in the following paragraphs.

For in-solution trypsinization, a reducing and denaturing buffer (50 mM TrisHCl, 10 mM DTT, 8 M urea) was added to 5 µg of purified GlycoDelete ASP1 in PBS, samples were incubated at 37°C for 45 minutes, a 10x concentrated iodoacetic acid solution (200 mM in nanopure H₂O) was added and samples incubated in the dark for another 30 minutes after which samples were concentrated via 0.5 ml spin columns (10 kDa molecular weight cutoff, VS0102; Sartorius) to 50 µl. To ensure maximal digestion the samples were treated with a mixture of trypsin and the endoprotease Lys-C (200 ng, 1:25 ratio of protease to target protein, Trypsin/Lys-C mix Mass Spec Grade; Promega). Samples were incubated 4 hours at 37°C to allow digestion by the Lys-C which is functional under highly denaturing conditions. Subsequently, the samples were diluted (450 µl 50 mM TrisHCl, pH 8), lowering the ureum concentration to activate the trypsin, incubated overnight at 37°C and dried in a speedvac. Samples were finally reconstituted with 10 µL of 0.2% trifluoroacetic acid (TFA) (Sigma-Aldrich) and cleaned up with C18 ZipTip® pipette tips (Millipore, Billerica, MA, USA) according to the manufacturer's

instructions. Samples were analyzed with α -cyano-4-hydroxycinnamic acid (CHCA) matrix saturated in 50% acetonitrile containing 0.1% TFA, on a 4800 MALDI TOF/TOFTM Analyzer (Applied Biosystems by Life Technologies) in the positive ion mode. The reported m/z values were observed in several iterations of technical optimization of these experiments and the results of the fully optimized experiments are shown. The m/z values observed in the mass spectra were compared to the theoretical m/z values of ASP1 tryptic peptides that were calculated using the PeptideMass software⁸. Values for different glycoforms of the tryptic glycopeptides were generated by the addition of glycan mass increments to those of the naked glycopeptides and values for ASP1 peptides resulting from nonspecific cleavage were calculated using the FindPept software⁹.

For in-gel tryptic digest, 8 μ g purified GlycoDelete ASP1 was mixed with 5x concentrated SDS-PAGE sample buffer (De Buck et al., 2012) and incubated for 8 minutes at 98°C. After the denaturation, an excess of iodoacetamide was added (3 μ l of 0.5 M iodoacetamide in nanopure H₂O), samples were incubated for one hour in the dark at room temperature, run on polyacrylamide gels and visualized via Coomassie staining. An area containing the GlycoDelete ASP1 carrying truncated glycans and possible larger glycoforms was excised, destained and dried in a speed vac. Subsequent tryptic digest was performed by reswelling the gel fragments with 500 ng trypsin / cLys-r mix resuspended in 25 mM ammonium hydrogen carbonate (NH₄HCO₃) and incubating overnight at 37°C. The tryptic peptides were extracted from the gel, dried in a speed vac, reconstituted with 10 μ L of 0.2% trifluoroacetic acid (TFA) (Sigma-Aldrich), cleaned up with C18 ZipTip® pipette tips and further analyzed as described above.

Immunization of rabbits

Two groups of five rabbits were immunized subcutaneously with purified GlycoDelete or wild type ASP1 (in LPS-free PBS) following a prime-boost protocol. Rabbits received a first injection of 35 μ g recombinant ASP1 mixed with complete Freund's adjuvant in a 1:1 ratio at day one, followed by three booster shots of 35 μ g ASP1 mixed with incomplete Freund's adjuvant in a 1:1 ratio at days 14, 28 and 56. Blood samples were taken on day zero (preimmune) and at days 28, 36 and 80 when the rabbits were euthanized. Serum samples were incubated for 30 minutes at 56°C prior to freezing. The trial was performed by the Health Department of the CER groupe (Marloie, Belgium).

⁸ http://web.expasy.org/peptide_mass/

⁹ <http://web.expasy.org/findpept/>

Characterization of the antibody response in rabbits

ELISAs were performed to study the antibodies induced by immunization of rabbits with the wild type and GlycoDelete ASP1. To measure the presence of IgG antibodies specific for the ASP1 peptide backbone, high-mannose bearing *cgI* ASP1 was reduced, denatured and deglycosylated with PNGaseF, and used as antigen in ELISA. To achieve the latter, PNGaseF buffer (50 mM Phosphate buffer, pH 7, 0.5% w/v SDS, 0.4 M β -mercapthoethanol) was added to purified *cgI* ASP1, samples were heated for 10 minutes at 98°C, Igepal was added (final concentration 1% w/v) together with in-house produced PNGaseF and samples were incubated overnight at 37°C. Complete ASP1 deglycosylation was monitored by mobility shift assay on Coomassie-stained polyacrylamide gels. A final 10 minute heating step at 98°C inactivated the PNGaseF. Reduced, denatured and deglycosylated ASP1 was then diluted in sodium carbonate buffer (50 mM, pH 9.6) and used to coat Maxisorp 96-well plates (Nunc, Sigma Aldrich) overnight at 4°C (10 ng per well). Wells were washed three times with washing buffer (PBS, 0.1% v/v Tween 20) and this step was repeated between all following steps unless stated differently. Wells were blocked for two hours (PBS, 250 mM glycine, 1% w/v BSA), emptied and allowed to dry overnight at room temperature. Serial dilution of sera were made (in PBS, 0.1% v/v Tween 20, 0.1% goat serum), 100 μ l was added to each well and plates were incubated for two hours. Donkey α -rabbit IgG-HRP (Sigma-Aldrich) was used as secondary antibody (diluted 1:5000 in PBS with 1% w/v BSA), 100 μ l was added and plates incubated for one hour. Plates were washed five times and 100 μ l TMB was added to allow color development that was stopped after 30 minutes by addition of 100 μ l 1M HCl after which plates were read at 450 nm in a VERSAmax tunable microplate reader. The background signal was subtracted from the measured values and an OD value of 0.3 was taken to calculate antibody titers.

A similar strategy was followed to probe for the presence of plant-glycan specific IgG antibodies in the sera of the immunized rabbits. Here, different glycoforms of GBP-Fc, a synthetic antibody, were used as antigen in ELISA to identify sugar specific total IgG. High-mannose and complex plant-glycan bearing glycoforms of the GBP-Fc had previously been purified from *Pichia pastoris* and wild type *Arabidopsis thaliana*, respectively, and had been extensively characterized (Dr. T. De Meyer T, PhD dissertation 2013). The naked GBP-Fc peptide backbone was obtained by reduction, denaturation and deglycosylation of the high-mannose GBP-Fc glycoform as described above for the *cgI* ASP1 and the peptide backbone was used together with the high-mannose and complex-glycan bearing glycoforms as antigen in ELISA. To avoid experimental variation due to differential antigen preparation, the high-mannose and complex glycoform were also treated as described above with the exception that heat-inactivated PNGaseF was added to the samples and this was checked by mobility shift assay on Coomassie-stained polyacrylamide gels. ELISA was further performed as described in the previous

paragraph with the exception that 50 ng of GBP-Fc was coated per well. The background signal, determined by coating the GBP-Fc peptide backbone as antigen, was subtracted from the measured values and an OD value of 0.3 was taken to calculate antibody titers.

Supplemental information

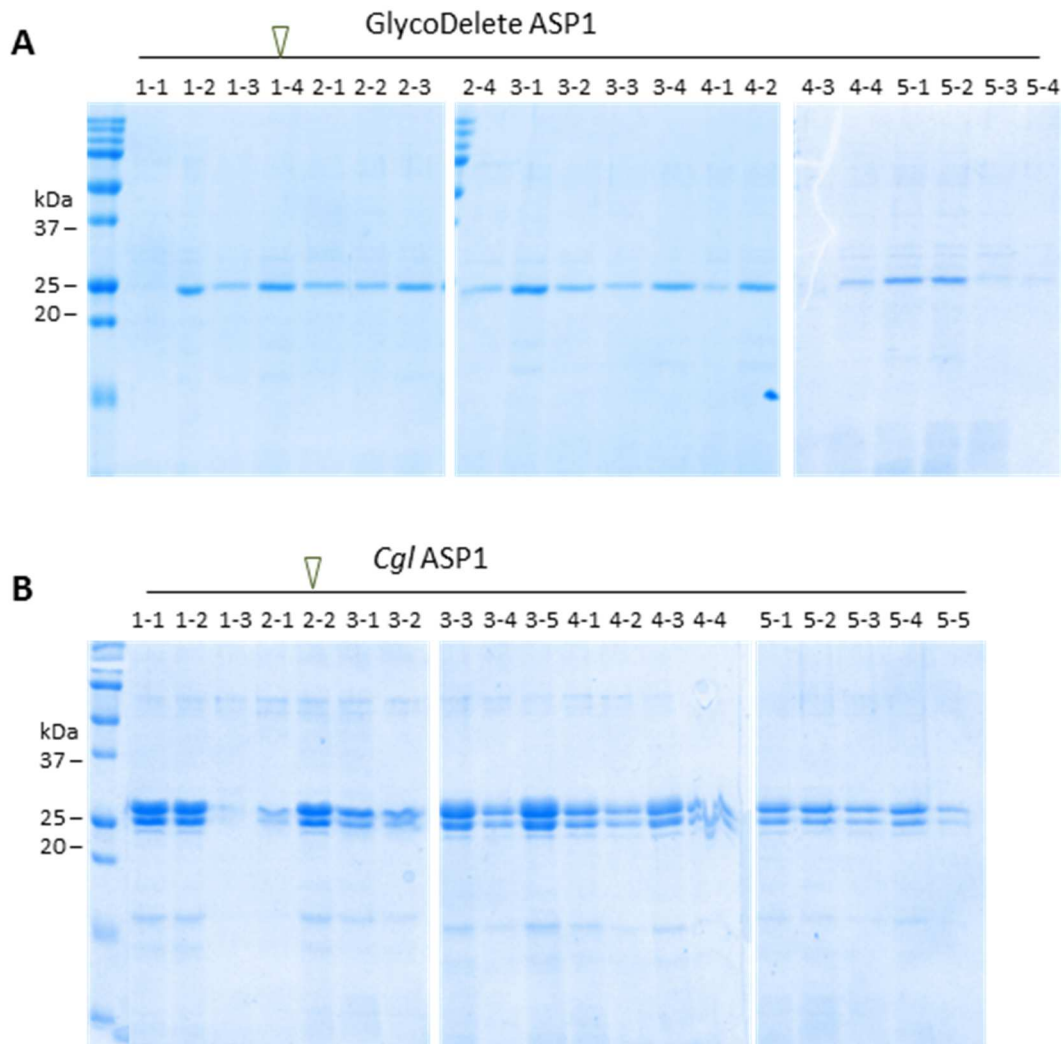


Figure S1. T2 analysis and inter-transformant variation in accumulation level of the seed-produced ASP1s. The wild type ASP1 was propagated to homozygosity by Dr. A. De Paepe and provided as such. Information on inter-transformant variation was not at hand for these lines. 20 GlycoDelete ASP1 and 19 *cgl* ASP1 T2 transformed seed stocks were screened for antigen content via Coomassie stained polyacrylamide gels. After visual inspection, seven lines showing high antigen accumulation were selected and analyzed for single locus T-DNA integration. Subsequently, the line with the highest antigen content and the ASP1 T-DNA integrated at a single locus in the genome, was retained (indicated with a hollow arrowhead) and propagated to homozygosity. Seed extracts were prepared at a fixed w/v ratio, 10 μ l was loaded onto the gels that were run under reducing conditions. (A) The 20 screened GlycoDelete ASP1 T2 seed stocks. (B) The 19 screened *cgl* ASP1 T2 seed stocks.

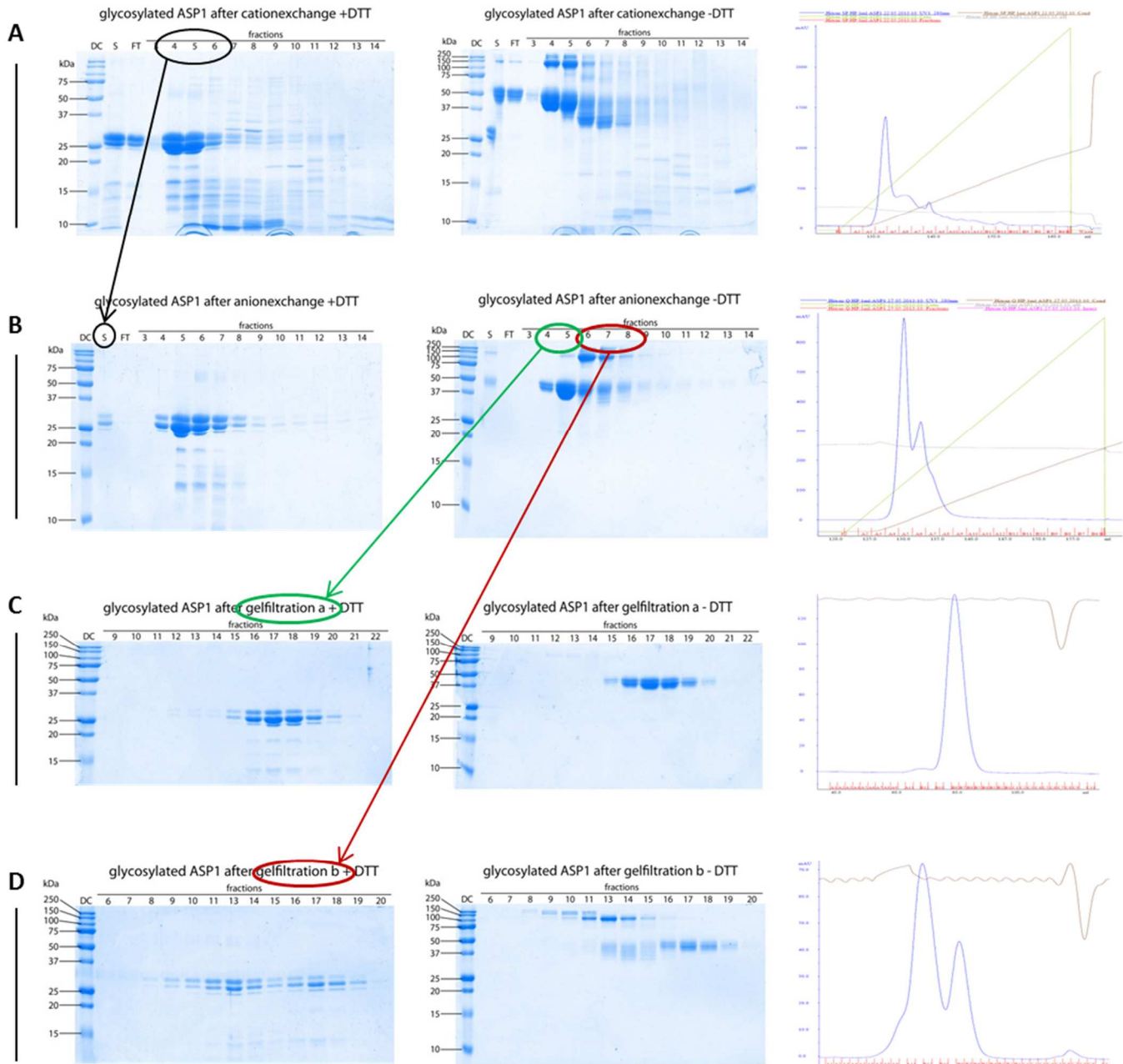


Figure S2. Recombinant ASP1 purification. Wild type ASP1 was extracted from one gram of a homozygous seed stock and purified by three consecutive chromatography steps. ASP1 purity and integrity was monitored by SDS-PAGE and Coomassie staining under reducing (left panels) and non-reducing conditions (middle panels). 10 μ l was loaded onto the gels. Right panels represent the corresponding chromatograms showing the OD280. (A) Cation exchange chromatography step. Fractions pooled, diluted and subjected to further purification are indicated with a black circle. (B) Anion exchange chromatography step. Eluted fractions were separated into ASP1 dimers (green circle) and multimers (red circle) and independently gel filtrated, (C) and (D), respectively. Abbreviations used: DC, Dual Color ladder; kDa, kilodalton; S, sample; FT, flow through.

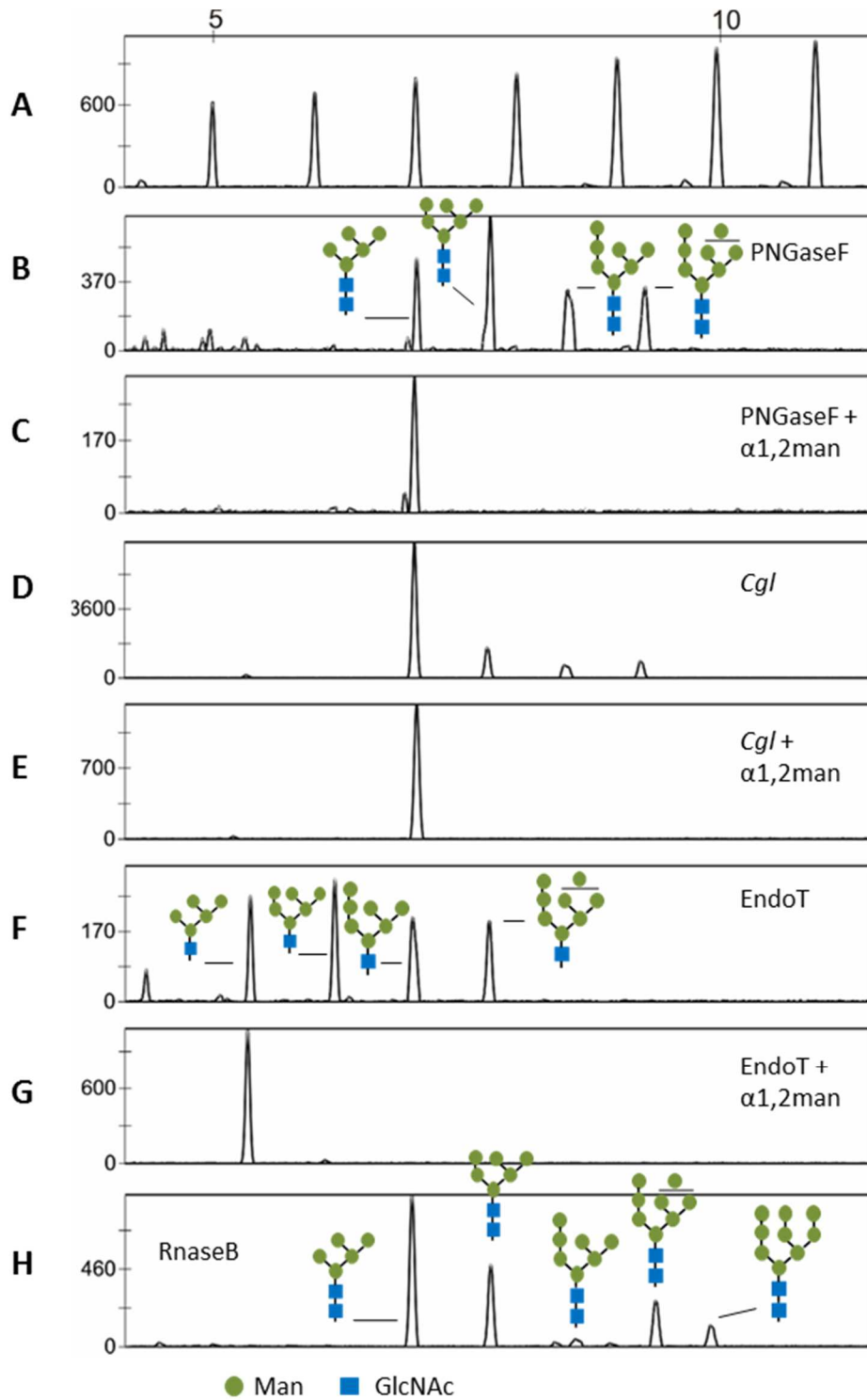


Figure S3. DSA-FACE analysis of recombinant ASP1 glycan structures and $\alpha 1,2$ -mannosidase ($\alpha 1,2$ man) digests. Glycans were released from purified recombinant ASP1 and treated with $\alpha 1,2$ -mannosidase. (A) Oligomaltose standard, the number of glucose units is indicated. (B) Glycans released from wild type ASP1 by PNGaseF and (C) digested with $\alpha 1,2$ man. (D) Glycans released from *cgl* ASP1 by PNGaseA and (E) digested with $\alpha 1,2$ man. (F) EndoT released wild type ASP1 glycans and (G) digested with $\alpha 1,2$ man. (H) RnaseB high-mannose glycan standard. Abbreviations used: $\alpha 1,2$ man, $\alpha 1,2$ -mannosidase.

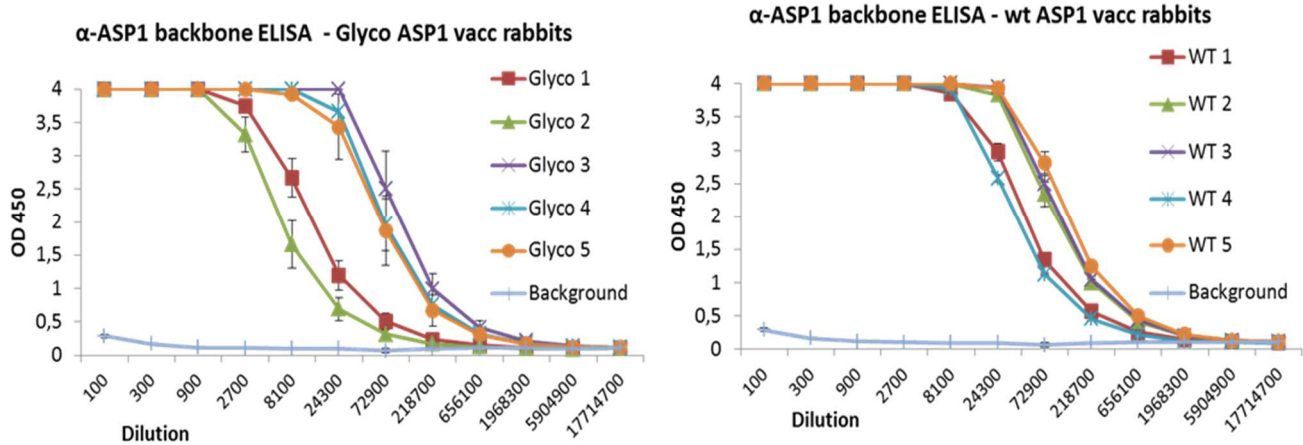


Figure S4. Overview of the α-ASP1 peptide backbone ELISAs. Sera obtained from the final bleeding of GlycoDelete (left panel) and wild type ASP1 immunized rabbits (right panel) were subjected to ELISA with reduced, denatured and deglycosylated ASP1 coated as the antigen. Values are the average of three independent experiments, error bars represent the standard deviation. Background indicates the signal detected when a pooled sample was loaded onto wells coated with inactivated PNGaseF only. Are given, the serum dilution and the corresponding signal measured at 450 nm.

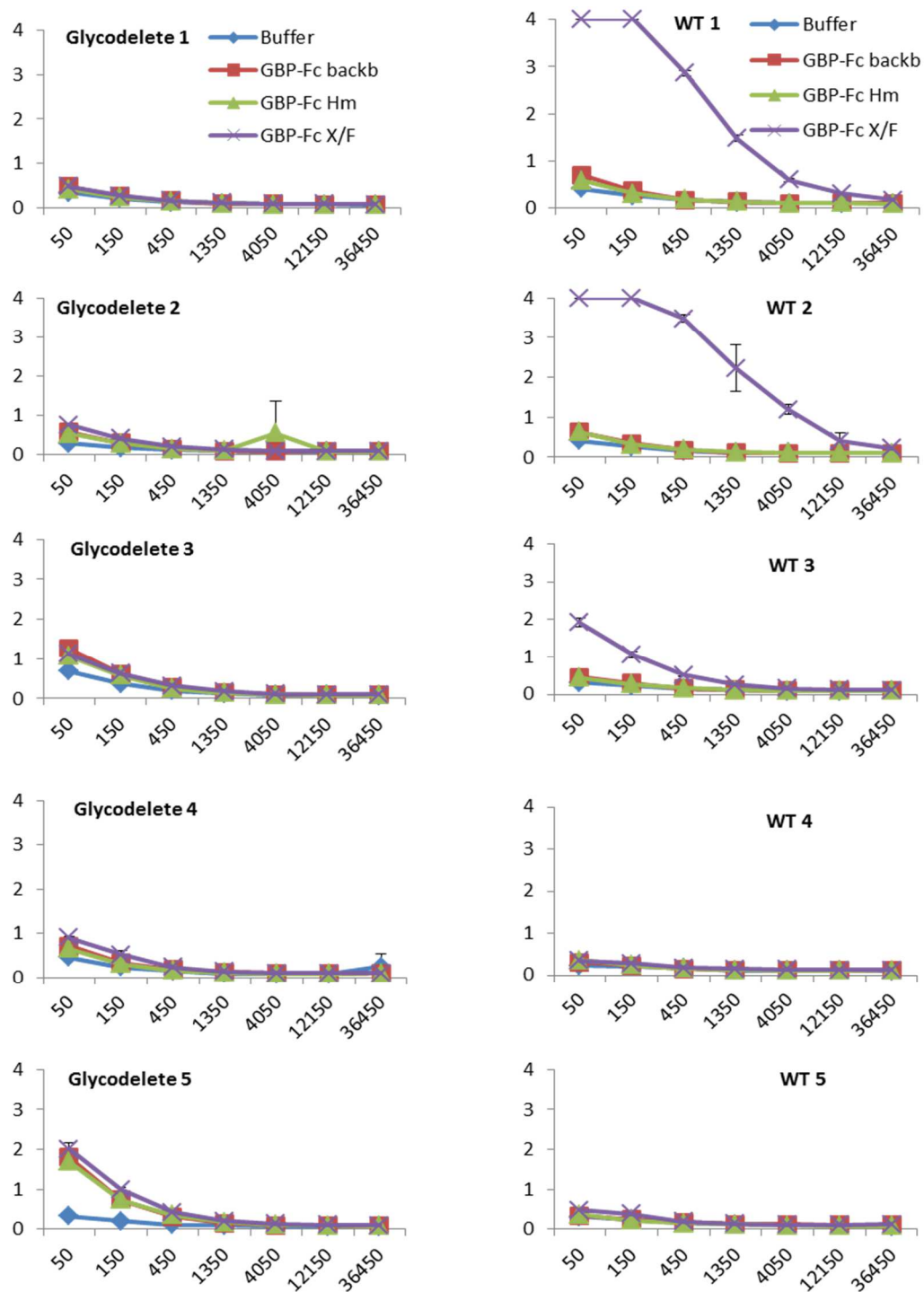


Figure S5. Overview of the α -GBP-Fc ELISAs. Sera obtained from the final bleeding of GlycoDelete (left panels) and wild type ASP1 immunized rabbits (right panels) were subjected to ELISA with the GBP-Fc peptide backbone, the high-mannose (Hm) GBP-Fc and the xylose/fucose (X/F) bearing GBP-Fc coated as antigens. The GBP-Fc peptide backbone was coated to determine the background signal. Are given, the serum dilution (X-axis) and the corresponding signal measured at 450 nm (Y-axis). Dilution series were made in triplicate and the average values are shown. No glycan-specific IgG could be detected in the sera of the GlycoDelete ASP1 immunized rabbits, whereas three out of five wild type ASP1 immunized rabbits had generated IgG specific for fucose and xylose.

Table S1. Matching peptides for unspecific cleavage of GlycoDelete ASP1

Observed m/z	Predicted m/z	Δ mass	peptide	position
1.975.020	1.974.967	-0.053	(L)RNLQPDTEANIYNWQI(R)	116-131
1.975.020	1.974.967	-0.053	(R)NLQPDTEANIYNWQIR(P)	117-132
1.975.020	1.974.967	-0.053	(S)LRNLQPDTEANIYNWQ(I)	115-130
1.975.020	1.974.992	-0.028	(I)DTSIPQNLAQWLLFQNS(Q)	81-97
1.975.020	1.975.083	0.063	(F)HNQVRRDIAGASPLLNT(G)	22-39
2.185.150	2.185.067	-0.082	(T)PWSWVTASLRNLQPDTEAN(I)	107-125
2.185.150	2.185.104	-0.046	(R)NLQPDTEANIYNWQIRPL(S)	117-134
2.185.150	2.185.129	-0.021	(L)PIDTSIPQNLAQWLLFQNS(Q)	79-97
2.386.220	2.386.179	-0.041	(R)NLQPDTEANIYNWQIRPLSN(I)	117-136

Supplemental table 1. *In silico* prediction of internal glycoDelete ASP1 peptides derived from nonspecific cleavage that match the m/z values of unidentified peaks in the MALDI-TOF spectrum of in-gel trypsinized GlycoDelete ASP1.

```

1   ATGAGTAAAC   GGAATCCGAA   GATTCTGAAG   ATTTTCTGT   ATATGTTACT
    TCTCAACTCT CTCTTTCTCA TCATCTACTT CGTTTTTCAC TCATCGGTAC
101 CCGTTAAAGA   ACTGCAGCTC   AGAGCTGAAC   CAACCGATCT   CCCTAGGCTT
    ATCGTCTACT TCCAGACTAC CCACGACTCT AGCAACCGTC CCATAAGCAT
201 GTTGCCACTC   ATCACTGAGA   AGGGTATAGC   CTTGACACAC   CTTATCGTTT
    GCTCCTTCCA CATTAACCAG GCGGAGTGG TTCACCTAAA CGACTTTCCA
301 CCCGATGATC   CACACTTCTA   CACCCTCTGG   AACGAGACTA   TCACTATGAA
    GCAGGCTGGG GTCAAGGTTA TGGGAATGGT CGGTGGCGCA GCTCCTGGTT
401 CATTCAACAC   GCAGACTCTT   GACTCTCCGG   ACTCAGCCAC   CTTCGAACAT
    TACTACGGCC AACTCCGCGA TGCTATAGTT AACTTCCAGT TGGAGGGAAT
501 GGATCTCGAC   GTGGAGCAAC   CAATGAGCCA   ACAAGGTATC   GACAGACTCA
    TCGCGCGACT AAGAGCTGAT TTCGACCCG ACTTCTTGAT TACCCTCGCC
601 CCTGTTGCTT   CAGCATTGGA   GGACTCCTCA   AACCTTTCTG   GTTTTAGCTA
    TACAGCCCTC CAACAGACCC AGGGAAACGA CATTGACTGG TACAACACTC
701 AGTTCTACTC   TGGTTTCGGC   TCAATGGCAG   ACACCTCCGA   CTACGACCGC
    ATCGTCGCTA ACGGCTTCGC CCCAGCTAAA GTTGTGCGCG GACAGCTCAC
801 GACTCCTGAA   GGTGCAGGAT   GGATTCCAAC   CAGCTCCCTA   AACAACACTA
    TCGTTAGCCT CGTATCCGAG TATGGCCAAA TAGGTGGGGT TATGGGATGG
901 GAGTACTTTA   ATAGCCTCCC   GGGTGAACCC   GCCGAACCAT   GGGAGTGGGC
    GCAAATCGTC ACAGAAATCC TTAGACCCGG ATTGGTTCCA GAGCTCAAGA
1001 TTACCGAGGA   TGACGCCGCT   CGTTTGACCG   GTGCATACGA   AGAAAGCGTG
    AAAGCTGCCG CTGCAGATAA CAAGTCCTTC GTTAAGAGGC CTAGCATCAA
1101 CTACTACGCT   ATGGTCAACG   CGTAA

```

```

1   MSKRNPKILK IFLYMLLLNS LFLIYFVFH SS VPKELQL RAEPTDLPRL IVYFQTTHDS

61  SNRPISMLPL ITEKGIALTH LIVCSFHINQ GGVVHLNDFP PDDPHFYTLW NETITMKQAG

121 VKVMGMVGGG APGSFNTQTL DSPDSATFEH YYGQLRDAIV NFQLEGMDLD VEQPMSQQGI

181 DRLIARLRAD FGPDLITLA PVASALEDSS NLSGFSYTAL QQTQGNDIDW YNTQFYSGFG

241 SMADTSDYDR IVANGFAPAK VVAGQLTTP E GAGWIPTSSL NNTIVSLVSE YGQIGGVMGW

301 EYFNSLPGGT AEPWEWAQIV TEILRPGLVP ELKITEDDAA RLTGAYEESV KAAAADNKSF

361 VKRPSINYA MVNA*

```

Supplemental Information 1. EndoT sequences. The EndoT open reading frame (ORF) from 5' to 3' end as it was cloned into the expression vector (upper sequence) and the translation thereof (lower sequence, the β 1,2-xylosyltransferase *medial*-Golgi targeting signal is underlined).

```

1   ATGGCAAACA  AGCTTTTCCT  CGTCTGCGCA  ACTTTCGCCC  TCTGCTTCCT
    CCTCACCAAC GCCGGTTTTT  GCTGCCCAGC  CGATCTCAAC  CAGACTGATG
101  AGGCAAGAAA  GATCTTCCTC  GATTTTCACA  ACCAAGTTCG  CCGTGATATC
    GCAGGTGCCA  GCCCGTTGCT  CAACCTCACC  GGAGCTGTTC  AGATGCGAAA
201  CGTTCTCGGT  CCAGCTAAGA  ACATGTACAG  AATGGACTGG  GACTGCAACC
    TCGAAGCCAA  AGCAAAGGCC  ATGATCTGGC  CATGCACTAC  CCCTCTCCCA
301  ATCGACACGA  GCATTCCACA  GAACCTCGCT  CAGTGGCTAC  TTTTCCAAAA
    CAGCCAGGAA  AACGAGGTCT  TGACCCAAAC  GCCCTGGTCT  TGGGTCACCG
401  CATCACTAAG  AAATCTTCAA  CCTGATACAG  AGGCTAACAT  CTACAACCTG
    CAGATTAGAC  CACTATCCAA  CATCGCTAAC  TGGCAGAACC  TAAAAGTTGG
501  ATGTGCTCAC  AAGGTTTGCA  AATTCCCCAC  CGGGACAAAC  ATGGTTGTGT
    CTTGCGCTTA  CGGCGGCGAG  GTCCTCCAGG  ATAACGAAGT  TGTCTGGGAC
601  AAGGGACCAA  CTTGCATGTG  CAACGCTTAC  CCAACAGCT  TCTGCTGCAA
    CAATCTCTGT  GACACAATCG  CTGCTGCAAC  ACTTCGCAAC  CAGCCTTGCA
701  AGAGCACTTG  A

```

```

1   MANKLFLVCA TFALCFLLTN A GFCCPADLN QTDEARKIFL DFHNQVRRDI AGASPLLNL
61  GAVQMRNVLG PAKNMYRMDW DCNLEAKAKA MIWPCTTPLP IDTSIPQNLQ QWLLFQNSQE
121 NEVLTQTPWS WVTASLRNLQ PDTEANIYNW QIRPLSNIAN WQNLKVGCAH KVCKFPTGTN
181 MVVSCAYGGE VLQDNEVVWD KGPTCMCNAY PNSFCCNNLC DTIAAATLRN QPCKST*

```

Supplemental information 2. Recombinant ASP1 sequences. The ASP1 ORF from 5' to 3' end as it was cloned into the expression vector (upper sequence) and the translation thereof (lower sequence, the 2S ER-signal peptide is underlined).

Chapter V

Conclusions & perspectives.

R.P. wrote the chapter and A. D. and N.C. edited it.

Vaccine production in *Arabidopsis* seeds

Since its discovery the PRRSV has spread rapidly, becoming endemic in almost all pig-rearing countries around the world, and is nowadays considered to be the most important pathogen of swine. Whereas older strains would mainly be problematic for pregnant sows and piglets, the situation is nowadays further aggravated by the fact that lately also highly pathogenic strains have been appearing that cause severe symptoms among pigs of all ages and have a high incidence of mortality. Vaccines are being used to minimize the clinical impact of an infection but because of the genetic drift of the virus, these constantly lose their power and it has been proposed that a whole battery of complementary, adaptable vaccines will be needed to effectively combat the virus. A subunit vaccine, that is safe and contains epitopes important in protective immunity, would be of good use in combination with the existing attenuated vaccines.

We wanted to evaluate the capacity of the *Arabidopsis* seed-platform, previously used to obtain high seed specific expression of a range of heterologous proteins, for vaccine production. To this goal we selected the GP3, GP4 and GP5 of the PRRSV to be expressed in our system and included in a subunit vaccine against the PRRSV. Viral antigens are complex, often difficult to express proteins and plants could occupy a niche among the current platforms by producing these. However, antigens are frequently plagued by low expressing levels that in their turn can obstruct further vaccine development. Indeed, when PRRSV antigens were expressed in different plant species they showed very low accumulation levels that did not even reach the 1% of TSP threshold. We set out to tackle this issue by the rational design of antigen formats and fusions that were aimed at augmenting the accumulation of the antigens in the seed.

Several strategies were tested and these included removal of transmembrane domains and/or coupling the antigens to stabilizing protein domains such as the GFP and IgG Fc-domains. We showed that the antigens as such do not accumulate very well in our system, hereby confirming their difficult to produce nature, but that by select optimization their accumulation could be boosted by multiple factors resulting in several antigens with yields that exceeded the 1% of TSP threshold, a feature that hasn't been attained by any previous study generating PRRSV antigens *in planta*. The results further show that the viral antigens can correctly be produced in the ER, with the additional benefit that high-mannose glycans are present on the antigens that might direct these to antigen presenting cells, and that the added Fc-domains have a double role and also effectively serve as affinity handle besides increasing accumulation.

We delivered proof-of-principle of the immunogenicity of purified seed-produced antigens by showing that virus-neutralizing antibodies are generated in mice after injection of a GP4 derived antigen, which

is similar to the *in vivo* situation where GP4 is also an important target for neutralizing antibodies. Alas, this could not be reproduced during a preliminary trial in piglets that had received an antigen cocktail. An in-depth experiment with challenge is required to shed light on this matter and to dissect the role of each antigen in more detail and evaluate if the antigens have, for example, a priming effect that stimulates a faster immune response after challenge. Perhaps the presence of an immunodominant, non-neutralizing B-cell epitope in the antigen cocktail skewed the induction of antibodies towards non-neutralizing epitopes. Indeed, for the GP5 such a decoy epitope has been identified (Ostrowski et al., 2002) and this might have affected the antibody response when it was administered together with the GP3 and GP4 derived antigens. Again, a challenge experiment could assess the role of each antigen alone and determine which antigens offer good protective immunity.

Another interesting venue is offered by the Fc-domain that is carried by the antigens that were administered to the piglets and that is derived from the porcine IgG3. *In silico* prediction hints at a good binding of the Fc-domain to the neonatal Fc receptor (FcRn) and experimental evidence seems to support this statement (Butler et al., 2009; Viridi et al., 2013). An alternative route of antigen administration could thus be envisaged where the antigens would be delivered intranasally and associate with the FcRns present in the mucosa via binding of their IgG3 Fc-domain, resulting in their effective uptake. The feasibility of this approach to mucosal vaccination was shown in mice by Ye et al. (2011) who could generate both a mucosal as a systemic B and T-cell response against a model antigen fused to an IgG Fc-domain when this was administered intranasally together with the adjuvant CpG. As the PRRSV is primarily transmitted via the respiratory route, a mucosal immune response would be a great first line of defence, which is a feature desirable for all PRRSV vaccines, and this merits further investigation.

All in all, our results confirm that the PRRSV antigens are poor accumulators *in planta*, and show that by mimicking the structures that do produce well in our system, i.e. antibodies, via fusion of the antigens to Fc-domains (and the GFP) their accumulation could be considerably increased. For some, but not all antigens, yields exceeded the 1% of TSP deemed necessary for economic viability of a PMP. The results also show that no general optimization rule fits all antigens, however, most antigens do have the tendency to accumulate better after fusion to stabilizing domains. Antigens could be efficiently purified and were immunogenic, at least in mice, but their protective capacity in pigs remains to be elucidated.

To formulate a straightforward conclusion concerning the suitability of the seed-platform for the production of antigens with the final intention of generating a functional vaccine, is a more complex matter. The obtained results show that, with respect to accumulation and purification, antigens can

indeed be produced in the seeds. The optimized antigens could be purified via a straightforward protocol and taking into consideration that downstream processing represents the bulk cost associated with therapeutic protein production, this turns out favourable. However, the protective capacity of the seed-produced antigens remains to be elucidated and this is as crucial as high yield and efficient purification. It determines factors such as the dose that is required per immunization and this in its turn directly influences the amount of antigen that needs to be produced and hence, the price tag associated with the final vaccine. To conclude, antigen production in the seed and purification from the latter is feasible, as shown by our results, but this not necessarily translates into an efficacious and affordable vaccine and the suitability of the seed platform for the production of such a vaccine must thus be evaluated on a case-by-case basis and will be determined by the outcome of the interplay between the factors antigen immunogenicity, accumulation and downstream processing.

Glycoengineering *Arabidopsis* seeds

Plants have distinct advantages over other systems for the production of recombinant proteins: they are capable of complex protein synthesis, they do not harbor animal and human pathogens, production is easily scalable for most systems and possibly, also cheap. One major drawback however, is that the plant and human N-glycosylation pathways diverge during the later stages of glycan maturation in the Golgi apparatus and this results in the introduction of specific structures on complex plant glycans that are not encountered on human glycans and hence, are possibly immunogenic. Indeed, a large part of the human population has antibodies against the plant specific glycans β 1,2-xylose and core α 1,3-fucose (Bardor et al., 2003; Mari, 2002) and this can prove problematic when a therapeutic protein bearing these glycans is delivered into the blood stream. Binding of these glycan specific antibodies can result in rapid clearance of the therapeutic, but in the worst case scenario, this immunogenicity might provoke a life threatening anaphylactic shock.

Given the importance of these immunogenic glycans, a plethora of strategies have been employed to 'humanize' plant glycans, i.e. to avoid the synthesis of immunogenic glycans and to convert plant glycosylation pathways to human-like synthesis. These glycoengineering techniques range from relatively straightforward retention of a recombinant protein in the ER, to the more elaborate inactivation of endogenous plant glycosyltransferases, which has often been coupled to the concomitant introduction of heterologous transferases into the plant glycosylation machinery. Needless to say, these are tedious strategies that require the coordinated expression of several genes at the same time and at the right place. These endeavours are further complicated by the polyploidy of many crop species, requiring decade long breeding strategies to obtain homozygous lines for all knock-outs and transgenes. We, on the other hand, have developed a straightforward system that bypasses extensive manipulations of the glycosylation pathway by seed-specific expression of a single fungal gene in a selected *Arabidopsis thaliana* mutant lacking just one endogenous glycosyltransferase. The resulting so-called 'GlycoDelete' seed-platform permits the production of any therapeutic glycoprotein and delivers the latter as one glycoform, bearing N-glycans that are trimmed down to a single GlcNAc, completely devoid of immunogenic β 1,2-xylose and core α 1,3-fucose.

To create such a platform, we selected the endoT, based on previous successes in mammalian cells, which is an endoglycosidase from the fungus *Trichoderma reesei* that cleaves high-mannose and hybrid-type N-glycans between the two protein-proximal β 1,4-linked GlcNAc residues. We fused it to a plant Golgi glycosyltransferase targeting signal and introduced it in the *Arabidopsis cgl* mutant whose proteins carry high-mannose glycans only, which are an excellent substrate for the endoT. We show that upon seed-specific expression, active endoT is produced that can digest the glycans of both

endogenous seed proteins and those of a reporter protein that was also introduced in the *cgl* seeds under control of the same regulatory elements used to drive *endoT* expression. Although the *endoT* content of the seeds was rather low, it was sufficient to entirely process a ten-fold excess of the reporter protein ASP1. The yield of the recombinant ASP1, an immunogenic protein from the helminth *Ostertagia ostertagi*, produced in the GlycoDelete seeds, was considerably lower than that of the ASP1 produced in seeds lacking *endoT* but the exact cause of this observation requires further in-depth studies. It should be noted that the ASP1 expression cassette was introduced independently in the *cgl* mutant and the Glycodelete background (*cgl* mutant plus *endoT*). Consequently, the lower expression might well be due to a clonal variation. The ASP1 has a high vaccine potential to protect cattle from *O. ostertagi* and because of difficulties in isolating large quantities of ASP1 from *Ostertagia* itself, there is a great interest to produce recombinant ASP1. However, recombinant versions of the ASP1 seem to lack some essential structural element, rendering them non-protective. Indications are at hand that also the seed-produced ASP1s are structurally divergent from the native ASP1 (neutralizing antibodies induced by the native ASP1 do not recognize recombinant ASP1 from wild type seeds). Further investigation is needed to elucidate the precise nature of this critical element and to determine if the seed-produced ASP1s also lack this element. One way to easily obtain additional information on this topic would be to check if the antibodies present in the sera of the rabbits that have been immunized with the wild type and the GlycoDelete ASP1 and bind to the peptide backbone of the seed-produced ASP1, are also capable of recognizing the native ASP1, e.g. by competitive ELISA. In case that protective capacity could be demonstrated for the seed-produced ASP1s, a recombinant vaccine against *Ostertagia ostertagi* based on the seed-produced ASP1s can surely be envisaged, given that the high accumulation of the ASP1s in the seed, combined with the optimized purification protocol would much facilitate ASP1 production.

In-depth study of the glycans present on the different seed-produced ASP1s revealed that the glycan profiles of the recombinant ASP1s are in complete accordance with the expected profiles: complex, immunogenic plant glycans for the wild type ASP1, high-mannose glycans for the *cgl* ASP1 and trimmed glycans consisting of a single GlcNAc on the GlycoDelete ASP1. Rabbits were injected with wild type and GlycoDelete ASP1 to investigate the immunogenicity of the glycans present on the ASP1s. Indeed, the majority, but not all of the rabbits immunized with wild type ASP1 had generated plant glycan specific antibodies that were completely absent in the sera of the GlycoDelete ASP1 immunized group. These results confirm the immunogenic potential of the plant specific β 1,2-xylose and core α 1,3-fucose so acclaimed in literature and validate the efforts dedicated to 'humanizing' plant glycans. It also, at least in rabbits, points to genetic predisposition as an important factor in determining the outcome of exposure to these potentially immunogenic glycans. Last but not least, it

convincingly demonstrates that the GlycoDelete seed-platform can be used for the *in planta* production of therapeutic proteins with a homogeneous glycan profile lacking immunogenic plant glycans.

The GlycoDelete seed platform has several unique advantages. The seed specific expression of the endoT (and the target protein) avoids interference with normal plant development and bestows the platform with the additional benefits associated with seed production of PMPs, such as accumulation of the PMP in a small volume and stable environment and the possibility to store the unprocessed seeds for prolonged periods of time, which allows decoupling of PMP production and processing. The platform is simple and only requires supertransformation of the GlycoDelete line with a gene of interest under control of the same regulatory elements used to drive the endoT gene. The system is robust and was shown to completely digest a recombinant protein present in ten-fold excess over the endoT, which translates into a minimal processing capacity of 7 mg recombinant protein per gram dry seed weight (3.5% of TSP) and possibly much more. The final and perhaps most important advantage is that it fully accomplishes the goals that were set during its conception, namely that it produces a therapeutic protein bearing short glycans, important to maintain its native conformation, that are homogeneous and at the same time also completely lack the immunogenic β 1,2-xylose and α 1,3-fucose.

The GlycoDelete seed-platform could be used as such for the commercial production of PMPs. The small stature of the *Arabidopsis* host facilitates cultivation in greenhouses where it can be densely packed and contained. Methods for the automatic harvesting and sowing of *Arabidopsis* seeds are under investigation and it has been estimated that *Arabidopsis* grown in a greenhouse facility could generate somewhere from 180 to 240 grams of dry seeds per square meter each year (He et al., 2013; Loos et al., 2011a). Using its minimal capacity as a conservative estimate, this would correspond to 1.2 to 1.6 grams of endoT digested PMP produced in the GlycoDelete seed platform per square meter on a yearly basis. Alternatively, the GlycoDelete technology could be transferred to another crop, such as pea, whose seeds have an even higher protein content that can account for up to 40% of seed weight. Work is underway in our lab to establish if expression patterns of synthetic IgAs produced in *Arabidopsis* seeds using the seed-specific expression cassette can be translated to pea (Dr. V. Viridi, personal communication). If the expression cassette functions similarly in pea then it would be feasible to transfer the GlycoDelete technology as such to pea and although a mutant lacking GnTI activity is not available, such a mutant could be created via meganuclease (or similar) technology.

Another interesting venue would be to check if the endoT could be transiently expressed in, for example, *Nicotiana benthamiana* leaves, as has been tested for the bacterial PNGaseF that was able to

completely deglycosylate co-expressed reporter proteins (Mamedov et al., 2012; Mamedov & Yusibov, 2013). PNGaseF has a broader substrate specificity than endoT, but it will nevertheless be interesting to determine if the endoT can also function as an integral part of a transient expression system. The latter represents another potential application of the endoT, especially because of the current industrial interest in platforms based on transient production of PMPs. As becomes clear from this discussion, there are several valid options available to continue with the application of the GlycoDelete technology *in planta*.

In conclusion, we have designed and engineered a platform for the seed-specific production of therapeutic proteins *in planta* that carry minimal glycans which are necessary to maintain the overall protein fold but that at the same time are homogeneous and completely devoid of the immunogenic plant specific structures β 1,2-xylose and α 1,3-fucose. We demonstrate the full functionality of the system and prove that a minimally glycosylated version of a recombinant helminth antigen produced in the platform does not provoke an anti-glycan response upon immunization of rabbits whereas its recombinant counterpart produced in non-engineered, wild type seeds does, hereby confirming the immunogenicity of these plant glycan structure and validating the potential of the GlycoDelete seed-platform. The system has several advantages over other glycoengineered solutions to 'humanize' plant glycans and can be used as such for industrial applications. Alternatively, the GlycoDelete technology can be tailored to meet specific needs and there are several options for further application of the technology *in planta*.

References

- Aalberse RC, van Ree R (1996) Cross-reactive carbohydrate determinants. *Monogr Allergy* **32**: 78-83
- Albina E, Carrat C, Charley B (1998) Interferon-alpha response to swine arterivirus (PoAV), the porcine reproductive and respiratory syndrome virus. *J Interferon Cytokine Res* **18**: 485-490
- Altmann F (2007) The role of protein glycosylation in allergy. *Int Arch Allergy Immunol* **142**: 99-115
- Alvarez ML, Cardineau GA (2010) Prevention of bubonic and pneumonic plague using plant-derived vaccines. *Biotechnol Adv* **28**: 184-196
- An TQ, Tian ZJ, Xiao Y, Li R, Peng JM, Wei TC, Zhang Y, Zhou YJ, Tong GZ (2010) Origin of Highly Pathogenic Porcine Reproductive and Respiratory Syndrome Virus, China. *Emerg Infect Dis* **16**: 365-367
- Ansari IH, Kwon B, Osorio FA, Pattnaik AK (2006) Influence of N-linked glycosylation of porcine reproductive and respiratory syndrome virus GP5 on virus infectivity, antigenicity, and ability to induce neutralizing antibodies. *J Virol* **80**: 3994-4004
- Anthony RM, Nimmerjahn F, Ashline DJ, Reinhold VN, Paulson JC, Ravetch JV (2008) Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG fc. *Science* **320**: 373-376
- Aviezer D, Brill-Almon E, Shaaltiel Y, Hashmueli S, Bartfeld D, Mizrahi S, Liberman Y, Freeman A, Zimran A, Galun E (2009) A Plant-Derived Recombinant Human Glucocerebrosidase Enzyme-A Preclinical and Phase I Investigation. *PLoS ONE* **4**: e4792
- Bakker H, Bardor M, Molthoff JW, Gomord V, Elbers I, Stevens LH, Jordi W, Lommen A, Faye L, Lerouge P, Bosch D (2001) Galactose-extended glycans of antibodies produced by transgenic plants. *Proc Natl Acad Sci USA* **98**: 2899-2904
- Bakker H, Rouwendal GJ, Karnoup AS, Florack DE, Stoop GM, Helsper JP, van Ree R, van Die I, Bosch D (2006) An antibody produced in tobacco expressing a hybrid beta-1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc Natl Acad Sci USA* **103**: 7577-7582
- Bardor M, Faveeuw C, Fitchette AC, Gilbert D, Galas L, Trottein F, Faye L, Lerouge P (2003) Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology* **13**: 427-434
- Bazin H (2011) *Vaccination--A History: From Lady Montagu to Genetic Engineering*: John Libbey Eurotext.

Boothe J, Nykiforuk C, Shen Y, Zaplachinski S, Szarka S, Kuhlman P, Murray E, Morck D, Moloney MM (2010) Seed-based expression systems for plant molecular farming. *Plant Biotechnol J* **8**: 588-606

Borloo J, Geldhof P, Peelaers I, Van Meulder F, Ameloot P, Callewaert N, Vercruysse J, Claerebout E, Strelkov SV, Weeks SD (2013) Structure of Ostertagia ostertagi ASP-1: insights into disulfide-mediated cyclization and dimerization. *Acta Crystallogr Sect D-Biol Crystallogr* **69**: 493-503

Bosch D, Castilho A, Loos A, Schots A, Steinkellner H (2013) N-Glycosylation of Plant-produced Recombinant Proteins. *Curr Pharm Design* **19**: 5503-5512

Bosch D, Schots A (2010) Plant glycans: friend or foe in vaccine development? *Expert Rev Vaccines* **9**: 835-842

Bublin M, Radauer C, Wilson IB, Kraft D, Scheiner O, Breiteneder H, Hoffmann-Sommergruber K (2003) Cross-reactive N-glycans of Api g 5, a high molecular weight glycoprotein allergen from celery, are required for immunoglobulin E binding and activation of effector cells from allergic patients. *FASEB J* **17**: 1697-1699

Buck PM, Kumar S, Singh SK (2013) Consequences of glycan truncation on Fc structural integrity. *mAbs* **5**: 904-916

Buonaguro FM, Butler-Ransohoff JE (2010) PharmaPlant: the new frontier in vaccines. Forward. *Expert Rev Vaccines* **9**: 805-807

Butler JE, Wertz N, Deschacht N, Kacs Kovics I (2009) Porcine IgG: structure, genetics, and evolution. *Immunogenetics* **61**: 209-230

Campbell CT, Yarema KJ (2005) Large-scale approaches for glycobiology. *Genome Biol* **6**: 236

Castilho A, Gattinger P, Grass J, Jez J, Pabst M, Altmann F, Gorfer M, Strasser R, Steinkellner H (2011) N-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. *Glycobiology* **21**: 813-823

Castilho A, Strasser R, Stadlmann J, Grass J, Jez J, Gattinger P, Kunert R, Quendler H, Pabst M, Leonard R, Altmann F, Steinkellner H (2010) In planta protein sialylation through overexpression of the respective mammalian pathway. *J Biol Chem* **285**: 15923-15930

Chan HT, Chia MY, Pang VF, Jeng CR, Do YY, Huang PL (2013) Oral immunogenicity of porcine reproductive and respiratory syndrome virus antigen expressed in transgenic banana. *Plant Biotechnol J* **11**: 315-324

Chargelegue D, Vine ND, van Dolleweerd CJ, Drake PMW, Ma JKC (2000) A murine monoclonal antibody produced in transgenic plants with plant-specific glycans is not immunogenic in mice. *Transgenic Res* **9**: 187-194

Chen X, Liu J (2011) Generation and immunogenicity of transgenic potato expressing the GP5 protein of porcine reproductive and respiratory syndrome virus. *J Virol Methods* **173**: 153-158

Chen ZL, Schuler MA, Beachy RN (1986) Functional-Analysis of Regulatory Elements in a Plant Embryo-Specific Gene. *Proc Natl Acad Sci USA* **83**: 8560-8564

Chia M-Y, Hsiao S-H, Chan H-T, Do Y-Y, Huang P-L, Chang H-W, Tsai Y-C, Lin C-M, Pang VF, Jeng C-R (2011) Evaluation of the immunogenicity of a transgenic tobacco plant expressing the recombinant fusion protein of GP5 of porcine reproductive and respiratory syndrome virus and B subunit of *Escherichia coli* heat-labile enterotoxin in pigs. *Vet Immunol Immunopathol* **140**: 215-225

Chia M-Y, Hsiao S-H, Chan H-T, Do Y-Y, Huang PL, Chang HW, Tsai Y-C, Lin C-M, Pang VF, Jeng C-R (2010) Immunogenicity of recombinant GP5 protein of porcine reproductive and respiratory syndrome virus expressed in tobacco plant. *Vet Immunol Immunopathol* **135**: 234-242

Chung CH, Mirakhur B, Chan E, Le Q, Berlin J, Morse M, Murphy BA, Satinover SM, Hosen J, Mauro D, Slebos RJ, Zhou QW, Gold D, Hatley T, Hicklin DJ, Platts-Mills TAE (2008) Cetuximab-induced anaphylaxis and IgE specific for galactose- α -1,3-galactose. *N Engl J Med* **358**: 1109-1117

Collins JE, Benfield DA, Christianson WT, Harris L, Hennings JC, Shaw DP, Goyal SM, McCullough S, Morrison RB, Joo HS, Gorcyca D, Chladek D (1992) Isolation of swine infertility and respiratory syndrome virus (isolate ATCC VR-2332) in North America and experimental reproduction of the disease in gnotobiotic pigs. *J Vet Diagn Invest* **4**: 117-126

Costers S, Lefebvre DJ, Van Doorselaere J, Vanhee M, Delputte PL, Nauwynck HJ (2010) GP4 of porcine reproductive and respiratory syndrome virus contains a neutralizing epitope that is susceptible to immunoselection in vitro. *Arch Virol* **155**: 371-378

Cox KM, Sterling JD, Regan JT, Gasdaska JR, Frantz KK, Peele CG, Black A, Passmore D, Moldovan-Loomis C, Srinivasan M, Cuisson S, Cardarelli PM, Dickey LF (2006) Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat Biotechnol* **24**: 1591-1597

Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY (1995) Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* **20**: 448-455

Czajkowsky DM, Hu J, Shao Z, Pleass RJ (2012) Fc-fusion proteins: new developments and future perspectives. *EMBO Mol Med* **4**: 1015-1028

Das PB, Dinh PX, Ansari IH, de Lima M, Osorio FA, Pattnaik AK (2010) The Minor Envelope Glycoproteins GP2a and GP4 of Porcine Reproductive and Respiratory Syndrome Virus Interact with the Receptor CD163. *J Virol* **84**: 1731-1740

De Buck S, Nolf J, De Meyer T, Viridi V, De Wilde K, Van Lerberge E, Van Droogenbroeck B, Depicker A (2013) Fusion of an Fc chain to a VHH boosts the accumulation levels in *Arabidopsis* seeds. *Plant Biotechnol J* **11**: 1006-1016

De Buck S, Viridi V, De Meyer T, De Wilde K, Piron R, Nolf J, Van Lerberge E, De Paepe A, Depicker A (2012) Production of camel-like antibodies in plants. *Methods Mol Biol* **911**: 305-324

De Jaeger G, Scheffer S, Jacobs A, Zambre M, Zobell O, Goossens A, Depicker A, Angenon G (2002) Boosting heterologous protein production in transgenic dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. *Nat Biotechnol* **20**: 1265-1268

De Wilde K, De Buck S, Vanneste K, Depicker A (2013) Recombinant Antibody Production in Arabidopsis Seeds Triggers an Unfolded Protein Response. *Plant Physiol* **161**: 1021-1033

Delputte PL, Nauwynck HJ (2004) Porcine Arterivirus infection of alveolar macrophages is mediated by sialic acid on the virus. *J Virol* **78**: 8094-8101

Delputte PL, Vanderheijden N, Nauwynck HJ, Pensaert MB (2002) Involvement of the matrix protein in attachment of porcine reproductive and respiratory syndrome virus to a heparinlike receptor on porcine alveolar macrophages. *J Virol* **76**: 4312-4320

Diamond J (1997) *Guns, Germs and Steel: The Fates of Human Societies*, NY: WW Norton & Company.

Díaz I, Darwich L, Pappaterra G, Pujols J, Mateu E (2006) Different European-type vaccines against porcine reproductive and respiratory syndrome virus have different immunological properties and confer different protection to pigs. *Virology* **351**: 249-259

Díaz I, Pujols J, Ganges L, Gimeno M, Darwich L, Domingo M, Mateu E (2009) In silico prediction and ex vivo evaluation of potential T-cell epitopes in glycoproteins 4 and 5 and nucleocapsid protein of genotype-I (European) of porcine reproductive and respiratory syndrome virus. *Vaccine* **27**: 5603-5611

Downing WL, Galpin JD, Clemens S, Lauzon SM, Samuels AL, Pidkowich MS, Clarke LA, Kermode AR (2006) Synthesis of enzymatically active human alpha-L-iduronidase in Arabidopsis cgl (complex glycan-deficient) seeds. *Plant Biotechnol J* **4**: 169-181

Downing WL, Hu XK, Kermode AR (2007) Post-transcriptional factors are important for high-level expression of the human alpha-L-iduronidase gene in Arabidopsis cgl (complex-glycan-deficient) seeds. *Plant Sci* **172**: 327-334

Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, Mukherjee PK, Zeilinger S, Grigoriev IV, Kubicek CP (2011) Trichoderma: the genomics of opportunistic success. *Nat Rev Microbiol* **9**: 749-759

Duran JP, Climent I, Sarraseca J, Urniza A, Cortés E, Vela C, Casal JI (1997) Baculovirus expression of proteins of porcine reproductive and respiratory syndrome virus strain Olot/91. Involvement of ORF3 and ORF5 proteins in protection. *Virus Genes* **14**: 19-29

Engering AJ, Cella M, Fluitsma D, Brockhaus M, Hoefsmit ECM, Lanzavecchia A, Pieters J (1997) The mannose receptor functions as a high capacity and broad specificity antigen receptor in human dendritic cells. *Eur J Immunol* **27**: 2417-2425

Fang Y, Snijder EJ (2010) The PRRSV replicase: Exploring the multifunctionality of an intriguing set of nonstructural proteins. *Virus Res* **154**: 61-76

Faye L, Boulaflous A, Benchabane M, Gomord W, Michaud D (2005) Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* **23**: 1770-1778

Fenner F, Henderson DA, Arita I, Jezek Z, Ladnyi ID (1988) *Smallpox and its eradication*, Vol. 6: World Health Organization Geneva.

Fitchette-Laine AC, Gomord V, Cabanes M, Michalski JC, Saint Macary M, Foucher B, Cavelier B, Hawes C, Lerouge P, Faye L (1997) N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J* **12**: 1411-1417

Foetisch K, Westphal S, Lauer I, Retzek M, Altmann F, Kolarich D, Scheurer S, Vieths S (2003) Biological activity of IgE specific for cross-reactive carbohydrate determinants. *J Allergy Clin Immunol* **111**: 889-896

Forthal DN, Gach JS, Landucci G, Jez J, Strasser R, Kunert R, Steinkellner H (2010) Fc-Glycosylation Influences Fc gamma Receptor Binding and Cell-Mediated Anti-HIV Activity of Monoclonal Antibody 2G12. *J Immunol* **185**: 6876-6882

Frank J, Kaulfuerst-Soboll H, Rips S, Koiwa H, von Schaewen A (2008) Comparative Analyses of Arabidopsis complex glycan1 Mutants and Genetic Interaction with staurosporin and temperature sensitive3a. *Plant Physiol* **148**: 1354-1367

Frazer IH (2004) Prevention of cervical cancer through papillomavirus vaccination. *Nat Rev Immunol* **4**: 46-54

Fujiyama K, Sakai Y, Misaki R, Yanagihara I, Honda T, Anzai H, Seki T (2004) N-linked glycan structures of human lactoferrin produced by transgenic rice. *Biosci Biotechnol Biochem* **68**: 2565-2570

Geldhof MF, Vanhee M, Van Breedam W, Van Doorselaere J, Karniychuk UU, Nauwynck HJ (2012) Comparison of the efficacy of autogenous inactivated Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) vaccines with that of commercial vaccines against homologous and heterologous challenges. *BMC Vet Res* **8**: 182

Geldhof P, Claerebout E, Knox D, Vercauteren I, Looszova A, Vercruysse J (2002) Vaccination of calves against *Ostertagia ostertagi* with cysteine proteinase enriched protein fractions. *Parasite Immunol* **24**: 263-270

Geldhof P, Claerebout E, Knox DP, Agneessens J, Vercruysse J (2000) Proteinases released in vitro by the parasitic stages of the bovine abomasal nematode *Ostertagia ostertagi*. *Parasitology* **121**: 639-647

Geldhof P, Meyvis Y, Vercruysse J, Claerebout E (2008) Vaccine testing of a recombinant activation-associated secreted protein (ASP1) from *Ostertagia ostertagi*. *Parasite Immunol* **30**: 57-60

Geldhof P, Vercauteren I, Gevaert K, Staes A, Knox DP, Vandekerckhove J, Vercruysse J, Claerebout E (2003) Activation-associated secreted proteins are the most abundant antigens in a host protective fraction from *Ostertagia ostertagi*. *Mol Biochem Parasitol* **128**: 111-114

Geldhof P, Vercauteren I, Vercruysse J, Knox DP, Van den Broeck W, Claerebout E (2004) Validation of the protective *Ostertagia ostertagi* ES-thiol antigens with different adjuvantia. *Parasite Immunol* **26**: 37-43

Giorgi C, Franconi R, Rybicki EP (2010) Human papillomavirus vaccines in plants. *Expert Rev Vaccines* **9**: 913-924

Goding JW (1978) Use of staphylococcal protein A as an immunological reagent. *J Immunol Methods* **20**: 241-253

Gomez L, Chrispeels MJ (1994) Complementation of an *Arabidopsis-Thaliana* Mutant That Lacks Complex Asparagine-Linked Glycans with the Human Cdna-Encoding N-Acetylglucosaminyltransferase-I. *Proc Natl Acad Sci USA* **91**: 1829-1833

Gomord V, Denmat L-A, Fitchette-Lainé A-C, Satiat-Jeunemaitre B, Hawes C, Faye L (1997a) The C-terminal HDEL sequence is sufficient for retention of secretory proteins in the endoplasmic reticulum (ER) but promotes vacuolar targeting of proteins that escape the ER. *Plant J* **11**: 313-325

Gomord V, Denmat LA, FitchetteLaine AC, SatiatJeunemaitre B, Hawes C, Faye L (1997b) The C-terminal HDEL sequence is sufficient for retention of secretory proteins in the endoplasmic reticulum (ER) but promotes vacuolar targeting of proteins that escape the ER. *Plant J* **11**: 313-325

Gomord V, Faye L (2004) Posttranslational modification of therapeutic proteins in plants. *Curr Opin Plant Biol* **7**: 171-181

Gomord V, Fitchette AC, Menu-Bouaouiche L, Saint-Jore-Dupas C, Plasson C, Michaud D, Faye L (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol J* **8**: 564-587

Goossens A, Dillen W, De Clercq J, Van Montagu M, Angenon G (1999) The *arcelin-5* gene of *Phaseolus vulgaris* directs high seed-specific expression in transgenic *Phaseolus acutifolius* and Arabidopsis plants. *Plant Physiol* **120**: 1095-1104

Granell A, Fernandez-del-Carmen A, Orzaez D (2010) In planta production of plant-derived and non-plant-derived adjuvants. *Expert Rev Vaccines* **9**: 843-858

He X, Galpin JD, Tropak MB, Mahuran D, Haselhorst T, von Itzstein M, Kolarich D, Packer NH, Miao YS, Jiang LW, Grabowski GA, Clarke LA, Kermode AR (2012a) Production of active human glucocerebrosidase in seeds of Arabidopsis thaliana complex-glycan-deficient (cgl) plants. *Glycobiology* **22**: 492-503

He X, Haselhorst T, von Itzstein M, Kolarich D, Packer NH, Kermode AR (2012b) Influence of an ER-retention signal on the N-glycosylation of recombinant human alpha-L-iduronidase generated in seeds of Arabidopsis. *Plant Mol Biol Rep* **79**: 157-169

He X, Pierce O, Haselhorst T, von Itzstein M, Kolarich D, Packer NH, Gloster TM, Voadlo DJ, Qian Y, Brooks D, Kermode AR (2013) Characterization and downstream mannose phosphorylation of human recombinant alpha-L-iduronidase produced in Arabidopsis complex glycan-deficient (cgl) seeds. *Plant Biotechnol J* **11**: 1034-1043

Herrera-estrella L, Depicker A, Vanmontagu M, Schell J (1983) Expression of Chimaeric Genes Transferred into Plant-Cells Using a Ti-Plasmid-Derived Vector. *Nature* **303**: 209-213

Hiatt A, Cafferkey R, Bowdish K (1989) Production of Antibodies in Transgenic Plants. *Nature* **342**: 76-78

Hu J, Ni Y, Dryman BA, Meng XJ, Zhang C (2012) Immunogenicity study of plant-made oral subunit vaccine against porcine reproductive and respiratory syndrome virus (PRRSV). *Vaccine* **30**: 2068-2074

Huether CM, Lienhart O, Baur A, Stemmer C, Gorr G, Reski R, Decker EL (2005) Glyco-engineering of moss lacking plant-specific sugar residues. *Plant Biol* **7**: 292-299

Jacobs PP, Callewaert N (2009) N-glycosylation Engineering of Biopharmaceutical Expression Systems. *Curr Mol Med* **9**: 774-800

Jin C, Altmann F, Strasser R, Mach L, Schahs M, Kunert R, Rademacher T, Glossl J, Steinkellner H (2008a) A plant-derived human monoclonal antibody induces an anti-carbohydrate immune response in rabbits. *Glycobiology* **18**: 235-241

Jin C, Hantusch B, Hemmer W, Stadimann J, Altmann F (2008b) Affinity of IgE and IgG against cross-reactive carbohydrate determinants on plant and insect glycoproteins. *J Allergy Clin Immunol* **121**: 185-190

Johnson CR, Griggs TF, Gnanandarajah J, Murtaugh MP (2011) Novel structural protein in porcine reproductive and respiratory syndrome virus encoded by an alternative ORF5 present in all arteriviruses. *J Gen Virol* **92**: 1107-1116

Kaneko Y, Nimmerjahn F, Ravetch JV (2006) Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **313**: 670-673

Karimi M, Inzé D, Depicker A (2002) GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci* **7**: 193-195

Karniychuk UU, Geldhof M, Vanhee M, Van Doorsselaere J, Saveleva TA, Nauwynck HJ (2010) Pathogenesis and antigenic characterization of a new East European subtype 3 porcine reproductive and respiratory syndrome virus isolate. *BMC Vet Res* **6**: 30

Key KF, Guenette DK, Yoon K-J, Halbur PG, Toth TE, Meng XJ (2003) Development of a heteroduplex mobility assay to identify field isolates of porcine reproductive and respiratory syndrome virus with nucleotide sequences closely related to those of modified live-attenuated vaccines. *J Clin Microbiol* **41**: 2433-2439

Kimman TG, Cornelissen LA, Moormann RJ, Rebel JMJ, Stochofe-Zurwieden N (2009) Challenges for porcine reproductive and respiratory syndrome virus (PRRSV) vaccinology. *Vaccine* **27**: 3704-3718

Kirnbauer R, Booy F, Cheng N, Lowy DR, Schiller JT (1992) Papillomavirus L1 Major Capsid Protein Self-Assembles into Virus-Like Particles That Are Highly Immunogenic. *Proc Natl Acad Sci USA* **89**: 12180-12184

Koncz C, Schell J (1986) The promoter of T_L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol Gen Genet* **204**: 383-396

Koprivova A, Stemmer C, Altmann F, Hoffmann A, Kopriva S, Gorr G, Reski R, Decker EL (2004) Targeted knockouts of *Physcomitrella* lacking plantspecific immunogenic N-glycans. *Plant Biotechnol J* **2**: 517-523

Krebbbers E, Herdies L, De Clercq A, Seurinck J, Leemans J, Van Damme J, Segura M, Gheysen G, Van Montagu M, Vandekerckhove J (1988) Determination of the processing sites of an *Arabidopsis* 2S albumin and characterization of the complete gene family. *Plant Physiol* **87**: 859-866

Kurosaka A, Yano A, Itoh N, Kuroda Y, Nakagawa T, Kawasaki T (1991) The Structure of a Neural Specific Carbohydrate Epitope of Horseradish-Peroxidase Recognized by Anti-Horseradish Peroxidase Antiserum. *J Biol Chem* **266**: 4168-4172

Labarque GG, Nauwynck HJ, Van Reeth K, Pensaert MB (2000) Effect of cellular changes and onset of humoral immunity on the replication of porcine reproductive and respiratory syndrome virus in the lungs of pigs. *J Gen Virol* **81**: 1327-1334

Lambrecht BN, Kool M, Willart MAM, Hammad H (2009) Mechanism of action of clinically approved adjuvants. *Curr Opin Immunol* **21**: 23-29

Laroy W, Contreras R, Callewaert N (2006) Glycome mapping on DNA sequencing equipment. *Nat Protoc* **1**: 397-405

Latypov RF, Hogan S, Lau H, Gadgil H, Liu D (2012) Elucidation of acid-induced unfolding and aggregation of human immunoglobulin IgG1 and IgG2 Fc. *J Biol Chem* **287**: 1381-1396

Lee C, Yoo D (2006) The small envelope protein of porcine reproductive and respiratory syndrome virus possesses ion channel protein-like properties. *Virology* **355**: 30-43

Lentz EM, Segretin ME, Morgenfeld MM, Wirth SA, Santos MJD, Mozgovoj MV, Wigdorovitz A, Bravo-Almonacid FF (2010) High expression level of a foot and mouth disease virus epitope in tobacco transplastomic plants. *Planta* **231**: 387-395

Liebmingner E, Veit C, Pabst M, Batoux M, Zipfel C, Altmann F, Mach L, Strasser R (2011) beta-N-Acetylhexosaminidases HEXO1 and HEXO3 Are Responsible for the Formation of Paucimannosidic N-Glycans in *Arabidopsis thaliana*. *J Biol Chem* **286**: 10793-10802

Loos A, Van Droogenbroeck B, Hillmer S, Grass J, Kunert R, Cao J, Robinson DG, Depicker A, Steinkellner H (2011a) Production of monoclonal antibodies with a controlled N-glycosylation pattern in seeds of *Arabidopsis thaliana*. *Plant Biotechnol J* **9**: 179-192

Loos A, Van Droogenbroeck B, Hillmer S, Grass J, Pabst M, Castilho A, Kunert R, Liang M, Arcalis E, Robinson DG, Depicker A, Steinkellner H (2011b) Expression of antibody fragments with a controlled N-glycosylation pattern and induction of endoplasmic reticulum-derived vesicles in seeds of *Arabidopsis*. *Plant Physiol* **155**: 2036-2048

Lopez OJ, Osorio FA (2004) Role of neutralizing antibodies in PRRSV protective immunity. *Vet Immunol Immunopathol* **102**: 155-163

Ma JKC, Drake PMW, Christou P (2003) The production of recombinant pharmaceutical proteins in plants. *Nat Rev Genet* **4**: 794-805

Ma JKC, Hikmat BY, Wycoff K, Vine ND, Chargelegue D, Yu L, Hein MB, Lehner T (1998) Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nat Med* **4**: 601-606

Maley F, Trimble RB, Tarentino AL, Plummer TH, Jr. (1989) Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal Biochem* **180**: 195-204

Mamedov T, Ghosh A, Jones RM, Mett V, Farrance CE, Musiychuk K, Horsey A, Yusibov V (2012) Production of non-glycosylated recombinant proteins in *Nicotiana benthamiana* plants by co-expressing bacterial PNGase F. *Plant Biotechnol J* **10**: 773-782

Mamedov T, Yusibov V (2013) In vivo deglycosylation of recombinant proteins in plants by co-expression with bacterial PNGase F. *Bioengineered* **4**: 338-342

Mardassi H, Massie B, Dea S (1996) Intracellular synthesis, processing, and transport of proteins encoded by ORFs 5 to 7 of Porcine Reproductive and Respiratory Syndrome Virus. *Virology* **221**: 98-112

Mari A (2002) IgE to cross-reactive carbohydrate determinants: Analysis of the distribution and appraisal of the in vivo and in vitro reactivity. *Int Arch Allergy Immunol* **129**: 286-295

Mari A, Ooievaar-de Heer P, Scala E, Giani M, Pirrotta L, Zuidmeer L, Bethell D, van Ree R (2008) Evaluation by double-blind placebo-controlled oral challenge of the clinical relevance of IgE antibodies against plant glycans. *Allergy* **63**: 891-896

Marillonnet S, Giritch A, Gils M, Kandzia R, Klimyuk V, Gleba Y (2004) In planta engineering of viral RNA replicons: efficient assembly by recombination of DNA modules delivered by *Agrobacterium*. *Proc Natl Acad Sci USA* **101**: 6852-6857

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EGJ, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barabote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*) (vol 26, pg 553, 2008). *Nat Biotechnol* **26**: 1193-1193

Mcaleer WJ, Buynak EB, Maigetter RZ, Wampler DE, Miller WJ, Hilleman MR (1984) Human Hepatitis-B Vaccine from Recombinant Yeast. *Nature* **307**: 178-180

Meulenbergh JJ, van Nieuwstadt AP, van Essen-Zandbergen A, Langeveld JP (1997a) Posttranslational processing and identification of a neutralization domain of the GP4 protein encoded by ORF4 of Lelystad virus. *J Virol* **71**: 6061-6067

Meulenbergh JJM, Demeijer EJ, Moormann RJM (1993a) Subgenomic RNAs of Lelystad Virus Contain a Conserved Leader Body Junction Sequence. *J Gen Virol* **74**: 1697-1701

Meulenbergh JJM, Hulst MM, de Meijer EJ, Moonen PLJM, den Besten A, de Kluiver EP, Wensvoort G, Moormann RJM (1993b) Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. *Virology* **192**: 62-72

Meulenbergh JJM, Petersen-Den Besten A, De Kluiver EP, Moormann RJM, Schaaper WMM, Wensvoort G (1995) Characterization of proteins encoded by ORFs 2 to 7 of Lelystad virus. *Virology* **206**: 155-163

Meulenbergh JJM, Petersen den Besten A, de Kluiver E, van Nieuwstadt A, Wensvoort G, Moormann RJM (1997b) Molecular characterization of Lelystad virus. *Vet Microbiol* **55**: 197-202

Meulenbergh JJM, Petersen den Besten A (1996) Identification and characterization of a sixth structural protein of Lelystad virus: The glycoprotein GP(2) encoded by ORF2 is incorporated in virus particles. *Virology* **225**: 44-51

Meyvis Y, Callewaert N, Gevaert K, Timmerman E, Van Durme J, Schymkowitz J, Rousseau F, Vercruysse J, Claerebout E, Geldhof P (2008) Hybrid N-glycans on the host protective activation-associated secreted proteins of *Ostertagia ostertagi* and their importance in immunogenicity. *Mol Biochem Parasitol* **161**: 67-71

Meyvis Y, Geldhof P, Gevaert K, Timmerman E, Vercruysse J, Claerebout E (2007) Vaccination against *Ostertagia ostertagi* with subfractions of the protective ES-thiol fraction. *Vet Parasitol* **149**: 239-245

Mimura Y, Church S, Ghirlando R, Ashton PR, Dong S, Goodall M, Lund J, Jefferis R (2000) The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol* **37**: 697-706

Misinzo GM, Delputte PL, Nauwynck HJ (2008) Involvement of proteases in porcine reproductive and respiratory syndrome virus uncoating upon internalization in primary macrophages. *Vet Res* **39**: 55

Morandini F, Avesani L, Bortesi L, Van Droogenbroeck B, De Wilde K, Arcalis E, Bazzoni F, Santi L, Brozzetti A, Falorni A, Stoger E, Depicker A, Pezzotti M (2011) Non-food/feed seeds as biofactories for the high-yield production of recombinant pharmaceuticals. *Plant Biotechnol J* **9**: 911-921

Muntz K (1998) Deposition of storage proteins. *Plant Mol Biol* **38**: 77-99

Nagels B, Santens F, Weterings K, Van Damme EJM, Callewaert N (2011a) Improved sample preparation for CE-LIF analysis of plant N-glycans. *Electrophoresis* **32**: 3482-3490

Nagels B, Van Damme EJM, Callewaert N, Zabeau L, Tavernier J, Delanghe JR, Boets A, Castilho A, Weterings K (2012a) Biologically active, magnICON (R)-expressed EPO-Fc from stably transformed *Nicotiana benthamiana* plants presenting tetra-antennary N-glycan structures. *J Biotechnol* **160**: 242-250

Nagels B, Van Damme EJM, Pabst M, Callewaert N, Weterings K (2011b) Production of Complex Multiantennary N-Glycans in *Nicotiana benthamiana* Plants. *Plant Physiol* **155**: 1103-1112

Nagels B, Weterings K, Callewaert N, Van Damme EJM (2012b) Production of plant made pharmaceuticals: from plant host to functional protein. *Crit Rev Plant Sci* **31**: 148-180

Nam H-M, Chae K-S, Song Y-J, Lee N-H, Lee J-B, Park S-Y, Song C-S, Seo K-H, Kang S-M, Kim M-C, Choi I-S (2013) Immune responses in mice vaccinated with virus-like particles composed of the GP5 and M proteins of porcine reproductive and respiratory syndrome virus. *Arch Virol* **158**: 1275-1285

Nauwynck HJ, Duan X, Favoreel HW, Van Oostveldt P, Pensaert MB (1999) Entry of porcine reproductive and respiratory syndrome virus into porcine alveolar macrophages via receptor-mediated endocytosis. *J Gen Virol* **80**: 297-305

Nauwynck HJ, Van Gorp H, Vanhee M, Karniyuchuk U, Geldhof M, Cao A, Verbeeck M, Van Breedam W (2012) Micro-dissecting the pathogenesis and immune response of PRRSV infection paves the way for more efficient PRRSV vaccines. *Transbound Emerg Dis* **59**: 50-54

Neumann EJ, Kliebenstein JB, Johnson CD, Mabry JW, Bush EJ, Seitzinger AH, Green AL, Zimmerman JJ (2005) Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. *J Am Vet Med Assoc* **227**: 385-392

Nielsen HS, Oleksiewicz MB, Forsberg R, Stadejek T, Bøtner A, Storgaard T (2001) Reversion of a live porcine reproductive and respiratory syndrome virus vaccine investigated by parallel mutations. *J Gen Virol* **82**: 1263-1272

Nimmerjahn F, Ravetch JV (2007) The antiinflammatory activity of IgG: the intravenous IgG paradox. *J Exp Med* **204**: 11-15

Obembe OO, Popoola JO, Leelavathi S, Reddy SV (2011) Advances in plant molecular farming. *Biotechnol Adv* **29**: 210-222

Obregon P, Chargelegue D, Drake PMW, Prada A, Nuttall J, Frigerio L, Ma JK-C (2006) HIV-1 p24-immunoglobulin fusion molecule: a new strategy for plant-based protein production. *Plant Biotechnol J* **4**: 195-207

Opriessnig T, Halbur PG, Yoon K-J, Pogranichniy RM, Harmon KM, Evans R, Key KF, Pallares FJ, Thomas P, Meng XJ (2002) Comparison of molecular and biological characteristics of a modified live porcine reproductive and respiratory syndrome virus (PRRSV) vaccine (Ingelvac PRRS MLV), the parent strain of the vaccine (ATCC VR2332), ATCC-VR2385, and two recent field isolates of PRRSV. *J Virol* **76**: 11837-11844

Ostrowski M, Galeota JA, Jar AM, Platt KB, Osorio FA, Lopez OJ (2002) Identification of neutralizing and nonneutralizing epitopes in the porcine reproductive and respiratory syndrome virus GP5 ectodomain (vol 76, pg 4241, 2002). *J Virol* **76**: 6863-6863

Pagny S, Bouissonnie F, Sarkar M, Follet-Gueye ML, Driouich A, Schachter H, Faye L, Gomord V (2003) Structural requirements for Arabidopsis beta1,2-xylosyltransferase activity and targeting to the Golgi. *Plant J* **33**: 189-203

Palacpac NQ, Yoshida S, Sakai H, Kimura Y, Fujiyama K, Yoshida T, Seki T (1999) Stable expression of human beta1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc Natl Acad Sci USA* **96**: 4692-4697

Paul M, Ma JK-C (2010) Plant-made immunogens and effective delivery strategies. *Expert Rev Vaccines* **9**: 821-833

Pearse BR, Hebert DN (2010) Lectin chaperones help direct the maturation of glycoproteins in the endoplasmic reticulum. *Biochim Biophys Acta* **1803**: 684-693

Peckham GD, Bugos RC, Su WW (2006) Purification of GFP fusion proteins from transgenic plant cell cultures. *Protein Expr Purif* **49**: 183-189

Pelosi A, Shepherd R, Walmsley AM (2012) Delivery of plant-made vaccines and therapeutics. *Biotechnol Adv* **30**: 440-448

Penney CA, Thomas DR, Deen SS, Walmsley AM (2011) Plant-made vaccines in support of the Millennium Development Goals. *Plant Cell Rep* **30**: 789-798

Perry RD, Fetherston JD (1997) *Yersinia pestis* - Etiologic agent of plague. *Clin Microbiol Rev* **10**: 35-66

Petrucelli S, Otegui MS, Lareu F, Dinh OT, Fitchette AC, Circosta A, Rumbo M, Bardor M, Carcamo R, Gomord V, Beachy RN (2006) A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnol J* **4**: 511-527

Piron R, De Koker S, De Paepe A, Goossens J, Grooten J, Nauwynck H, Depicker A (2014) Boosting In Planta Production of Antigens Derived from the Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) and Subsequent Evaluation of Their Immunogenicity. *PLoS ONE* **9**: e91386

Plana-Durán J, Bastons M, Urniza A, Vayreda M, Vilà X, Mañé H (1997) Efficacy of an inactivated vaccine for prevention of reproductive failure induced by porcine reproductive and respiratory syndrome virus. *Vet Microbiol* **55**: 361-370

Plotkin SA (2005) Vaccines: past, present and future. *Nat Med* **11**: S5-11

Plotkin SA, Plotkin SL (2011) The development of vaccines: how the past led to the future. *Nat Rev Microbiol* **9**: 889-893

Prieto C, Martínez-Lobo FJ, Díez-Fuertes F, Aguilar-Calvo P, Simarro I, Castro JM (2011) Immunisation of pigs with a major envelope protein sub-unit vaccine against porcine reproductive and respiratory syndrome virus (PRRSV) results in enhanced clinical disease following experimental challenge. *Vet J* **189**: 323-329

Rappuoli R (2007) Bridging the knowledge gaps in vaccine design. *Nat Biotechnol* **25**: 1361-1366

Rappuoli R, Miller HI, Falkow S (2002) Medicine - The intangible value of vaccination. *Science* **297**: 937-939

Rinaudo CD, Telford JL, Rappuoli R, Seib KL (2009) Vaccinology in the genome era. *J Clin Invest* **119**: 2515-2525

Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC, Leonhardt H (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol Cell Proteomics* **7**: 282-289

Rouwendal GJ, Florack DE, Hesselink T, Cordewener JH, Helsper JP, Bosch D (2009) Synthesis of Lewis X epitopes on plant N-glycans. *Carbohydr Res* **344**: 1487-1493

Rouwendal GJA, Wuhler M, Florack DEA, Koeleman CAM, Deelder AM, Bakker H, Stoop GM, van Die I, Helsper JPDFG, Hokke CH, Bosch D (2007) Efficient introduction of a bisecting GlcNAc residue in tobacco N-glycans by expression of the gene encoding human N-acetylglucosaminyltransferase III. *Glycobiology* **17**: 334-344

Rowland RRR (2010) The interaction between PRRSV and the late gestation pig fetus. *Virus Res* **154**: 114-122

Rybicki EP (2010) Plant-made vaccines for humans and animals. *Plant Biotechnol J* **8**: 620-637

Rybicki ER (2009) Plant-produced vaccines: promise and reality. *Drug Discov Today* **14**: 16-24

Saint-Jore-Dupas C, Faye L, Gomord V (2007) From planta to pharma with glycosylation in the toolbox. *Trends Biotechnol* **25**: 317-323

Salyaev RK, Rigano MM, Rekoslavskaya NI (2010) Development of plant-based mucosal vaccines against widespread infectious diseases. *Expert Rev Vaccines* **9**: 937-946

Shchelkunov SN, Shchelkunova GA (2010) Plant-based vaccines against human hepatitis B virus. *Expert Rev Vaccines* **9**: 947-955

Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YG, Weikert SHA, Presta LG (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc gamma RIII and antibody-dependent cellular toxicity. *J Biol Chem* **277**: 26733-26740

Shields RL, Namenuk AK, Hong K, Meng YG, Rae J, Briggs J, Xie D, Lai J, Stadlen A, Li B, Fox JA, Presta LG (2001) High resolution mapping of the binding site on human IgG1 for Fc gamma RI, Fc gamma RII, Fc gamma RIII, and FcRn and design of IgG1 variants with improved binding to the Fc gamma R. *Int J Dev Biol* **276**: 6591-6604

Snijder EJ, Meulenbergh JJM (1998) The molecular biology of arteriviruses. *J Gen Virol* **79**: 961-979

Snijder EJ, van Tol H, Pedersen KW, Raamsman MJB, de Vries AAF (1999) Identification of a novel structural protein of arteriviruses. *J Virol* **73**: 6335-6345

Sourrouille C, Marquet-Blouin E, D'Aoust MA, Kiefer-Meyer MC, Seveno M, Pagny-Salehabadi S, Bardor M, Durambur G, Lerouge P, Vezina L, Gomord V (2008) Down-regulated expression of plant-specific glycoepitopes in alfalfa. *Plant Biotechnol J* **6**: 702-721

Stadejek T, Oleksiewicz MB, Scherbakov AV, Timina AM, Krabbe JS, Chabros K, Potapchuk D (2008) Definition of subtypes in the European genotype of porcine reproductive and respiratory syndrome virus: nucleocapsid characteristics and geographical distribution in Europe. *Arch Virol* **153**: 1479-1488

Stals I, Karkehabadi S, Kim S, Ward M, Van Landschoot A, Devreese B, Sandgren M (2012) High resolution crystal structure of the endo-N-acetyl- β -D-glucosaminidase responsible for the deglycosylation of *Hypocrea jecorina* cellulases. *PLoS ONE* **7**: e40854

Stals I, Samyn B, Sergeant K, White T, Hoorelbeke K, Coorevits A, Devreese B, Claeysens M, Piens K (2010) Identification of a gene coding for a deglycosylating enzyme in *Hypocrea jecorina*. *FEMS Microbiol Lett* **303**: 9-17

Stoger E, Ma JKC, Fischer R, Christou P (2005) Sowing the seeds of success: pharmaceutical proteins from plants. *Curr Opin Biotechnol* **16**: 167-173

Strasser R, Altmann F, Mach L, Glossl J, Steinkellner H (2004) Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett* **561**: 132-136

Strasser R, Castilho A, Stadlmann J, Kunert R, Quendler H, Gattinger P, Jez J, Rademacher T, Altmann F, Mach L, Steinkellner H (2009) Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous beta1,4-galactosylated N-glycan profile. *Int J Dev Biol* **284**: 20479-20485

Strasser R, Mucha J, Mach L, Altmann F, Wilson IB, Glossl J, Steinkellner H (2000) Molecular cloning and functional expression of beta1, 2-xylosyltransferase cDNA from *Arabidopsis thaliana*. *FEBS Lett* **472**: 105-108

Strasser R, Stadlmann J, Schahs M, Stiegler G, Quendler H, Mach L, Glossl J, Weterings K, Pabst M, Steinkellner H (2008) Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol J* **6**: 392-402

Strasser R, Stadlmann J, Svoboda B, Altmann F, Glossl J, Mach L (2005) Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex N-glycans. *Biochem J* **387**: 385-391

Sun Z, Chen ZH, Lawson SR, Fang Y (2010) The Cysteine Protease Domain of Porcine Reproductive and Respiratory Syndrome Virus Nonstructural Protein 2 Possesses Deubiquitinating and Interferon Antagonism Functions. *J Virol* **84**: 7832-7846

Tan MCAA, Mommaas AM, Drijfhout JW, Jordens R, Onderwater JJM, Verwoerd D, Mulder AA, van der Heiden AN, Scheidegger D, Oomen LCJM, Ottenhoff THM, Tulp A, Neefjes JJ, Koning F (1997) Mannose receptor-mediated uptake of antigens strongly enhances HLA class II-restricted antigen presentation by cultured dendritic cells. *Eur J Immunol* **27**: 2426-2435

Terpstra C, Wensvoort G, Pol JMA (1991) Experimental reproduction of porcine epidemic abortion and respiratory syndrome (mystery swine disease) by infection with Lelystad vims: Koch's postulates fulfilled. *Vet Q* **13**: 131-136

Thanavala Y, Mahoney M, Pal S, Scott A, Richter L, Natarajan N, Goodwin P, Arntzen CJ, Mason HS (2005) Immunogenicity in humans of an edible vaccine for hepatitis B. *Proc Natl Acad Sci USA* **102**: 3378-3382

Tian K, Yu X, Zhao T, Feng Y, Cao Z, Wang C, Hu Y, Chen X, Hu D, Tian X, Liu D, Zhang S, Deng X, Ding Y, Yang L, Zhang Y, Xiao H, Qiao M, Wang B, Hou L, Wang X, Yang X, Kang L, Sun M, Jin P, Wang S,

Kitamura Y, Yan J, Gao G (2007) Emergence of fatal PRRSV variants: unparalleled outbreaks of atypical PRRS in China and molecular dissection of the unique hallmark. *PLoS ONE* **2**: e526

Umana P, Jean-Mairet J, Moudry R, Amstutz H, Bailey JE (1999) Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat Biotechnol* **17**: 176-180

Van Breedam W, Costers S, Vanhee M, Gagnon CA, Rodríguez-Gómez IM, Geldhof M, Verbeeck M, Van Doorselaere J, Karniychuk U, Nauwynck HJ (2011) Porcine reproductive and respiratory syndrome virus (PRRSV)-specific mAbs: supporting diagnostics and providing new insights into the antigenic properties of the virus. *Vet Immunol Immunopathol* **141**: 246-257

Van Breedam W, Delputte PL, Van Gorp H, Misinzo G, Vanderheijden N, Duan XB, Nauwynck HJ (2010a) Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. *J Gen Virol* **91**: 1659-1667

Van Breedam W, Van Gorp H, Zhang JQ, Crocker PR, Delputte PL, Nauwynck HJ (2010b) The M/GP₅ glycoprotein complex of porcine reproductive and respiratory syndrome virus binds the sialoadhesin receptor in a sialic acid-dependent manner. *PLoS Pathog* **6**: e1000730

Van Coppennolle S, Van Meulder F, Mihi B, Peelaers I, Vercruysse J, Claerebout E, Geldhof P (2012) Analysis of the vaccine induced immune response against the abomasal parasite *Ostertagia ostertagi* in cattle suggests a pivotal role for NK cells. *Immunology* **137**: 755-755

Van Droogenbroeck B, Cao J, Stadlmann J, Altmann F, Colanesi S, Hillmer S, Robinson DG, Van Lerberge E, Terryn N, Van Montagu M, Liang M, Depicker A, De Jaeger G (2007) Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic *Arabidopsis* seeds. *Proc Natl Acad Sci USA* **104**: 1430-1435

Van Gorp H, Van Breedam W, Delputte PL, Nauwynck HJ (2008) Sialoadhesin and CD163 join forces during entry of the porcine reproductive and respiratory syndrome virus. *J Gen Virol* **89**: 2943-2953

van Ree R (2002) Carbohydrate epitopes and their relevance for the diagnosis and treatment of allergic diseases. *Int Arch Allergy Immunol* **129**: 189-197

Vanhee M, Costers S, Van Breedam W, Geldhof MF, Van Doorselaere J, Nauwynck HJ (2010) A variable region in GP4 of European-type porcine reproductive and respiratory syndrome virus induces neutralizing antibodies against homologous but not heterologous virus strains. *Viral Immunol* **23**: 403-413

Vanhee M, Delputte PL, Delrue I, Geldhof MF, Nauwynck HJ (2009) Development of an experimental inactivated PRRSV vaccine that induces virus-neutralizing antibodies. *Vet Res* **40**: 63

Vanhee M, Van Breedam W, Costers S, Geldhof M, Noppe Y, Nauwynck H (2011) Characterization of antigenic regions in the porcine reproductive and respiratory syndrome virus by the use of peptide-specific serum antibodies. *Vaccine* **29**: 4794-4804

Vashisht K, Goldberg TL, Husmann RJ, Schnitzlein W, Zuckermann FA (2008) Identification of immunodominant T-cell epitopes present in glycoprotein 5 of the North American genotype of porcine reproductive and respiratory syndrome virus. *Vaccine* **26**: 4747-4753

Vimolmangkang S, Gasic K, Soria-Guerra R, Rosales-Mendoza S, Moreno-Fierros L, Korban SS (2012) Expression of the nucleocapsid protein of porcine reproductive and respiratory syndrome virus in soybean seed yields an immunogenic antigenic protein. *Planta* **235**: 513-522

Virdi V, Coddens A, De Buck S, Millet S, Goddeeris BM, Cox E, De Greve H, Depicker A (2013) Orally fed seeds producing designer IgAs protect weaned piglets against enterotoxigenic *Escherichia coli* infection. *Proc Natl Acad Sci USA* **110**: 11809-11814

Virdi V, Depicker A (2013) Role of plant expression systems in antibody production for passive immunization. *Int J Dev Biol* **57**: 587-593

von Schaewen A, Frank J, Koiwa H (2008) Role of complex N-glycans in plant stress tolerance. *Plant Signal Behav* **3**: 871-873

von Schaewen A, Sturm A, O'Neill J, Chrispeels MJ (1993) Isolation of a mutant Arabidopsis plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol* **102**: 1109-1118

Vu HLX, Kwon B, Yoon KJ, Laegreid WW, Pattnaik AK, Osorio FA (2011) Immune Evasion of Porcine Reproductive and Respiratory Syndrome Virus through Glycan Shielding Involves both Glycoprotein 5 as Well as Glycoprotein 3. *J Virol* **85**: 5555-5564

Walsh G, Jefferis R (2006) Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* **24**: 1241-1252

Wang CC, Chen JR, Tseng YC, Hsu CH, Hung YF, Chen SW, Chen CM, Khoo KH, Cheng TJ, Cheng YSE, Jan JT, Wu CY, Ma C, Wong CH (2009) Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc Natl Acad Sci USA* **106**: 18137-18142

Wang W, Chen X, Xue C, Du Y, Lv L, Liu Q, Li X, Ma Y, Shen H, Cao Y (2012) Production and immunogenicity of chimeric virus-like particles containing porcine reproductive and respiratory syndrome virus GP5 protein. *Vaccine* **30**: 7072-7077

Wenderoth I, von Schaewen A (2000) Isolation and characterization of plant N-acetyl glucosaminyltransferase I (GntI) cDNA sequences. Functional analyses in the *Arabidopsis* cgl mutant and in antisense plants. *Plant Physiol* **123**: 1097-1108

Wensvoort G, Terpstra C, Pol JMA, ter Laak EA, Bloemraad M, de Kluyver EP, Kragten C, van Buiten L, den Besten A, Wagenaar F, Broekhuijsen JM, Moonen PLJM, Zetstra T, de Boer EA, Tibben HJ, de Jong MF, van 't Veld P, Greenland GJR, van Gennep JA, Voets MT, Verheijden JHM, Braamskamp J (1991) Mystery swine disease in the Netherlands: The isolation of Lelystad virus. *Vet Q* **13**: 121-130

Wissink EHJ, Kroese MV, van Wijk HAR, Rijsewijk FAM, Meulenbergh JJM, Rottier PJM (2005) Envelope protein requirements for the assembly of infectious virions of porcine reproductive and respiratory syndrome virus. *J Virol* **79**: 12495-12506

Wu W-H, Fang Y, Farwell R, Steffen-Bien M, Rowland RRR, Christopher-Hennings J, Nelson EA (2001) A 10-kDa structural protein of porcine reproductive and respiratory syndrome virus encoded by ORF2b. *Virology* **287**: 183-191

Wu W-H, Fang Y, Rowland RRR, Lawson SR, Christopher-Hennings J, Yoon K-J, Nelson EA (2005) The 2b protein as a minor structural component of PRRSV. *Virus Res* **114**: 177-181

Ye L, Zeng R, Bai Y, Roopenian DC, Zhu X (2011) Efficient mucosal vaccination mediated by the neonatal Fc receptor. *Nat Biotechnol* **29**: 158-163

Yusibov V, Streatfield SJ, Kushnir N (2011) Clinical development of plant-produced recombinant pharmaceuticals: Vaccines, antibodies and beyond. *Hum Vaccin* **7**: 313-321

Zeitlin L, Olmsted SS, Moench TR, Co MS, Martinell BJ, Paradkar VM, Russell DR, Queen C, Cone RA, Whaley KJ (1998) A humanized monoclonal antibody produced in transgenic plants for immunoprotection of the vagina against genital herpes. *Nat Biotechnol* **16**: 1361-1364

Zhou L, Yang H (2010) Porcine reproductive and respiratory syndrome in China. *Virus Res* **154**: 31-37

Zuckermann FA, Garcia EA, Luque ID, Christopher-Hennings J, Doster A, Brito M, Osorio F (2007) Assessment of the efficacy of commercial porcine reproductive and respiratory syndrome virus (PRRSV) vaccines based on measurement of serologic response, frequency of gamma-IFN-producing cells and virological parameters of protection upon challenge. *Vet Microbiol* **123**: 69-85

Addendum

Acknowledgements

A great many people have been involved in this work, all of which I would like to profoundly thank for their contribution, and several have been of critical importance for its successful completion. None more than my guiding professor Ann Depicker, who I met when I took one of her courses during the Master after Master program in Plant Biotechnology which I followed at the Ghent University. After having completed a Biochemistry career at the Catholic University of Leuven, filled with biophysics, enzyme kinetics and chemistry, but very little molecular biology, I was fascinated by the world of genetic engineering and realized that coming to Ghent, the cradle of plant biotechnology, had been a good choice. Motivated to pursue a career in plant biotechnology I then contacted Ann to inquire about the possibilities to start a PhD in her research group and yes indeed, a few conversations later I was delighted to hear that she would accept me as her PhD student (under the condition that I'd successfully apply for an IWT scholarship, but that's another story). However, one issue kept creeping up on me. Ever since reading the adventures of Tintin as a child, I had dreamt of one day venturing abroad to get to know far away countries and the people who lived there. Not only had I not fulfilled this dream, I was about to compromise myself to staying another four years in the land of chocolate, fries and beer... Eventually, I contacted Ann to tell her that despite really appreciating her offer (and having wasted her precious time), I had decided to travel to South America for some months instead. Not sure how she would respond, I received an email that I remember to the letter until today and that, to my astonishment, said the following: 'Dear Robin, you should not feel obliged under any societal pressure whatsoever to start a PhD; if you feel that at this point in time travelling will be more enriching for you as a person, then you should do so. If you would still be interested in a PhD position upon your return, feel free to contact me to discuss the possibilities. Ann.' And indeed, about a year later I followed her advice (again) and started in her group as a fresh and well-travelled PhD student. Need I say more? Thank you Ann, for the patience, for the opportunities you have given me and for always supporting me.

Let me extend my gratefulness to the research groups of Prof. Dr. Nauwynck, Prof. Dr. Callewaert and Prof. Dr. Grooten for their contribution to the practical work described in this thesis, and also our own research group, both the current lab members Jonah, Vikram, Els, Thomas and Veronique as our former members Sylvie, Kirsten, Annelies, Annick, Jolien and Joke. Thank you for the fun, nice lunch conversations and ofcourse, the scientific discussions and helping hands during the practical work. I would also like to thank my good friends Silke and Nathan for supporting me during the conception of

this thesis and keeping my mental health in check by organizing an occasional and unexpected evening out. An additional honorary mentioning is in place here for the latter, who also served as a sparring partner during martial arts training.

Now that I've touched upon the social theme, there are quite some more people that have colored my life here in Ghent during the years I've spent working as a PhD student. I'll kick off with my youth friends from hometown Antwerp; great friends, great party! Then those that I've met here in Ghent, in the lab, outside of the lab, during climbing, caving or related activities: Linde, Wannes & Wannes, Hans, Lorin, François, my Spanish speaking friends, and many, many more, just too many to mention. Thank you for being present in my life! I'd also like to express my exceptional gratitude to François Van Damme who, besides being a friend, willingly designed the cover of my thesis.

Then there is this very special group of people without whom I would simply not be where I stand now. I ofcourse refer to my family, my parents Nico and Ans, my sister Elena and brother Florian that have always been there for me and supported me during whichever endeavor I embarked on: from going to live alone in Leuven over choosing adventure above a PhD back to encouraging me to apply for a PhD scholarship a year later and even watch me participate in martial arts competition. Thank you, all of you (also my uncles, aunts, cousins and grandparents), for your unquestionable support.

Finally, there is one person that merits a paragraph dedicated only to her: Patricia Alvarado, whom I met during my sabbatical year in South America and who has been my girlfriend, companion, partner in crime and the woman I love ever since. It is said that 'behind every great man stands a strong woman', well, I am trying to live up to the expectations... Thank you for the five years we have shared and thank you for your unconditional love.

Curriculum Vitae

Personal information

First & last name	Robin Piron
Gender:	Male
Date & place of birth:	18-09-1985, Wilrijk, Belgium
Nationality:	Belgian
Social security number:	85091834590



Contact details

Permanent address:	Joos Robijnslei 54, 2140, Borgerhout, Belgium
Email address:	Robin.piron@hotmail.com , Pironrsh@gmail.com
Cellular phone no.:	0032 474 81 70 03
Phone no.:	0032 3 322 89 40

Education

PhD (2009 – 2014): Ghent University, Faculty of Sciences, Plant-made antibodies and immunogens research group (Prof. Dr. A. Depicker). Title of the PhD dissertation: 'Vaccine production and Glycoengineering in *Arabidopsis thaliana* seeds'.

Successfully completed the Doctoral Training Programme (soft skills, seminars and specialist courses) organized by the Doctoral School of Life Sciences and Medicine.

Master after Master in Plant Biotechnology (2007 – 2008):

Ghent University, Faculty of Sciences, Title of the dissertation: 'Hormonal crosstalk between the auxin and cytokinin signalling pathway in the root meristem of *Arabidopsis thaliana*'. Obtained with great distinction.

Master in Biochemistry (2005 – 2007):

Catholic University of Leuven, Faculty of Sciences, Title of the dissertation: 'Biosynthese van steviol, het aglycon van de zoetstof stevioside, in *Stevia rebaudiana* Bertoni'. Obtained with distinction.

Bachelor in Biology (2003 – 2005):

Catholic University of Leuven, Faculty of Sciences. Obtained with distinction.

Training sessions attended

Laboratory animal courses I & II, Ghent University (Certified for working with Lab animals in Belgium, 2011)

Effective scientific communication course, by Dr. Jean-luc Doumount, Ghent University (2011)

Tech transfer & Science based entrepreneurship course, VIB Research Training Course (2011)

Communication skills: conflict, negotiation and networking skills, Ghent University (2011)

Nanobody and protein production training, by Prof. Dr. S Muyldermans, VIB Research Training Course, (2010)

Scientific activities

Attended the 'Plant-based vaccines, antibodies and biologics' (PBVA) conference in Verona, Italy, 5-7 June 2013.

Participated from 15 August to 15 September 2012 in field work for the study 'Diversidad de tres tipos de simbioses (parasitos sanguineos, acaros de las plumas y virus causante de lesiones cutaneas) de la avifauna Del bosque protector El Bosque, Loja, Ecuador' in the Podocarpus National Park, Ecuador, supervised by biologist M. Moens.

Attended the COST (European Cooperation in Science and Technology) molecular farming action (FA0804) meeting in Vienna, Austria, 16-17 February 2012.

Attended the COST (European Cooperation in Science and Technology) molecular farming action (FA0804) meeting in Ghent, Belgium, 14-16 September 2011.

Language skills

Dutch	Mother tongue
English	In-depth reading, writing and conversational skills
Spanish	In-depth reading and conversational skills, good writing skills
French	Good reading and conversational skills, basic writing skills

Personal interests

Travelling, outdoor activities, development cooperation and cultural studies.

Publications

Piron R, De Koker S, De Paepe A, Goossens J, Grooten J, Nauwynck H, Depicker A (2014) Boosting In Planta Production of Antigens Derived from the Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) and Subsequent Evaluation of Their Immunogenicity. *PLoS ONE* **9**: e91386

De Buck S, Viridi V, De Meyer T, De Wilde K, Piron R, Nolf J, Van Lerberge E, De Paepe A, Depicker A (2012) Production of camel-like antibodies in plants. *Methods Mol Biol* **911**: 305-324